

STARS

University of Central Florida
STARS

Electronic Theses and Dissertations, 2004-2019

2014

Instructional Strategies for Scenario-Based Training of Human Behavior Cue Analysis with Robot-Aided Intelligence, Surveillance, Reconnaissance

Julie Salcedo
University of Central Florida



Part of the [Engineering Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Salcedo, Julie, "Instructional Strategies for Scenario-Based Training of Human Behavior Cue Analysis with Robot-Aided Intelligence, Surveillance, Reconnaissance" (2014). *Electronic Theses and Dissertations, 2004-2019*. 1304.

<https://stars.library.ucf.edu/etd/1304>



INSTRUCTIONAL STRATEGIES FOR SCENARIO-BASED TRAINING OF
HUMAN BEHAVIOR CUE ANALYSIS WITH ROBOT-AIDED
INTELLIGENCE, SURVEILLANCE, RECONNAISSANCE

by

JULIE NANETTE SALCEDO
B.S. University of Central Florida, 2005
M.S. University of Central Florida, 2011

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Modeling and Simulation
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2014

Major Professor: Stephanie J. Lackey

© Julie Nanette Salcedo

ABSTRACT

The U.S. Army desires to improve safety during Intelligence, Surveillance, Reconnaissance (ISR) operations by removing Warfighters from direct line-of-fire by enhancing ISR operational capabilities with unmanned systems, also known as Robot-Aided ISR (RAISR) (DOD, 2013). Additionally, RAISR presents an opportunity to fulfill ISR capability requirements of modern combat environments including: detection of High-Value Individuals (HVI) from safer distances, identification of baseline behavior, and interpretation of adversarial intent (U.S. Army, 2008). Along with the demand and projected acquisition of RAISR technology, there is the added need to design training requirements for system operation and task execution instruction. While documentation identifying specific training standards and objectives for ISR tasks utilizing unmanned systems is limited (DOD, 2013), simulation-based training has been identified as a critical training medium for RAISR (U.S. Army, 2008). ISR analysts will primarily conduct RAISR tasks via Indirect Vision Displays (IVD) which transition well into multimodal simulations (Salcedo, Lackey, & Maraj, 2014). However, simulation alone may not fulfill the complex training needs of RAISR tasks, therefore, incorporating instructional support may improve the effectiveness of training (Oser, Gualtieri, Cannon-Bowers, & Salas, 1999). One method to accomplish this is to utilize a Scenario-Based Training (SBT) framework enhanced with instructional strategies to target specific training objectives.

The purpose for the present experiment was to assess the effectiveness of SBT enhanced with selected instructional strategies for a PC-based RAISR training simulation. The specific task type was the identification of HVIs within a group through behavior cue analysis. The instructional strategies assessed in this experiment, Highlighting and Massed Exposure, have

shown to improve attentional weighting, visual search, and pattern recognition skills, which are critical for successful behavior cue analysis. Training effectiveness was evaluated by analyzing the impact of the instructional strategies on performance outcomes, including detection accuracy, classification accuracy, and median response time, and perceptions of the level of engagement, immersion, and presence during training exercises. Performance results revealed that the Massed Exposure strategy produced significantly faster response times for one subtle and one familiar target behavior cue. Perception results indicated that Highlighting was the least challenging instructional strategy and the Control offered the preferred level of challenge. The relationships between performance and perception measures revealed that higher levels of engagement, immersion, and presence were associated with better performance in the Control, but this trend did not always hold for Massed Exposure and Highlighting. Furthermore, presence emerged as the primary predictor of performance for select target behavior cues in the Control and Massed Exposure conditions, while immersion and engagement predicted performance of select cues in the Highlighting condition. The findings of the present experiment point to the potential benefit of SBT instructional strategies to improve effectiveness of simulation-based training for behavior cue analysis during RAISR operations. Specifically, the findings suggest that the Massed Exposure strategy has the potential to improve response time when detecting both familiar and novel targets. The results also highlight directions for future research to investigate methods to alter instructional strategy design and delivery in order to improve trainee perceptions of the instruction.

Dedicated with love and appreciation to my encouraging and supportive husband and parents.

ACKNOWLEDGMENTS

Foremost, to the One from whom all blessings flow, I am grateful that my eyes were opened and my steps were guided toward this path where the potential of my talents have been realized. I recognize that this achievement is only a stepping stone along my journey and I humbly anticipate the opportunities and challenges that lie ahead.

To my husband, Bernard, thank you for always encouraging me to shine, for patiently enduring the moments of my “absence,” for celebrating the milestones, and for expressing genuine pride in my achievements. You are the love of my life and I am continually blessed by your loyal and loving support.

To my parents, Sam and Nancy, I am grateful to have parents with such complementary qualities. Dad, you are my model of persistence, hard work, loyalty, and scholarship. Mom, you are my model of creative thinking, faithfulness, resourcefulness, and humility. Together, you have cultivated in me a boldness to think outside the box, a passion to pursue knowledge and truth, an appreciation of both failure and achievement, and a desire to help and encourage others in their pursuits. Thank you for your wise counsel and gentle guidance. You both are a testament to the verse from Proverbs 22:6, “Train up a child in the way she should go, and when she is grown she will not depart from it.”

To my sister, Nyki, I am thankful to have a kind, tender, loving sibling with whom I can rejoice in life’s blessings. As enthusiastic cheerleaders, we have rooted for each other in our shared and individual endeavors. Thank you for your continued support and sisterly pride.

To all my faithful family and supportive friends, I thank you for all your prayers and words of encouragement. Your kindness and concern did not go unnoticed.

To the faculty and students of the UCF ACTIVE Lab, thank you for the roles you each played in supporting the various aspects of this experiment from scenario development to data collection to technical support to document proofreading. I do not stand alone in my success.

To the Army Research Laboratory Human Research Directorate (ARL-HRED), I am grateful for the opportunity provided to make a meaningful contribution that will benefit the service men and women devoted to defend our freedoms.

Finally, to my future family, I pray that you regard this achievement, not as a burden you must live up to, but as an example of recognizing your talents and diligently pursuing goals that tap the full potential of your abilities.

TABLE OF CONTENTS

LIST OF FIGURES	xii
LIST OF TABLES	xiv
LIST OF ACRONYMS/ABBREVIATIONS	xv
CHAPTER ONE: INTRODUCTION	1
CHAPTER TWO: LITERATURE REVIEW	4
Intelligence, Surveillance, Reconnaissance	4
Intelligence.....	4
Surveillance.....	5
Reconnaissance	6
ISR Perceptual Skills	6
Robot-Aided ISR	7
RAISR Perceptual Skills.....	8
High-Value Individuals.....	9
Combat Profiling.....	11
Role of Combat Profiling in ISR	15
Training.....	20
Scenario-Based Training.....	20
Scenario-Based Training: Adaptive, Intelligent, Dynamic	23
Instructional Strategies.....	25
Training Effectiveness Evaluation	28

Performance	29
Engagement, Immersion, and Presence	30
Purpose of Present Experiment.....	32
CHAPTER THREE: METHODOLOGY	34
Participants.....	34
Experimental Testbed	35
Experimental Design.....	40
Independent Variable	42
Control Condition	42
Massed Exposure Condition	42
Highlighting Condition	42
Dependent Variables	43
Performance Metrics	43
Questionnaires.....	44
Procedure	46
CHAPTER FOUR: RESULTS	51
Outliers.....	51
Data Transformations.....	51
Analyses of Performance Metrics	52
Detection Accuracy.....	52
Classification Accuracy	53

Median Response Time.....	54
Percent Change	57
Analyses of Questionnaires.....	60
Engagement Measure.....	60
Immersion Measure	61
Presence Measure.....	62
Correlates of Performance Metrics and Questionnaires	63
Control Group Correlates.....	63
Massed Exposure Group Correlates.....	69
Highlighting Group Correlates	72
Regressions	73
Control Group Regressions.....	73
Massed Exposure Group Regressions.....	74
Highlighting Group Regressions.....	75
CHAPTER FIVE: DISCUSSION.....	76
CHAPTER SIX: CONCLUSION	82
APPENDIX A: ENGAGEMENT MEASURE.....	86
APPENDIX B: IMMERSION MEASURE	88
APPENDIX C: PRESENCE MEASURE	90
APPENDIX D: INFORMED CONSENT FORM	94
APPENDIX E: DEMOGRAPHICS QUESTIONNAIRE	98
APPENDIX F: EXAMPLES OF TARGET BEHAVIOR CUE MODELS	101

APPENDIX G: KINESIC CUE REVIEW SCORE SHEET	104
APPENDIX H: INSTITUTIONAL REVIEW BOARD APPROVAL LETTER.....	106
LIST OF REFERENCES	108

LIST OF FIGURES

Figure 1. Types of Human Terrain Analysis and Biometrics-Enabled Intelligence data.	10
Figure 2. Proactive approach with Combat Profiling versus reactive response to an attack. Adapted from (Ross & Militello, 2013).	12
Figure 3. The six domains of Combat Profiling.	15
Figure 4. Alignment of ISR data types for HVIs and the six Combat Profiling domains.	16
Figure 5. ISR task, strategy, tool hierarchy.	17
Figure 6. Components of the Scenario-Based Training framework. Adapted from (Fowlkes, Dwyer, Oser, & Salas, 1998; Lyons, Schmorrow, Cohn, & Lackey, 2002).	22
Figure 7. Components of Scenario-Based Training: Adaptive, Intelligent, Dynamic. Adapted from (Nicholson, Schatz, Stanney, & Lackey, 2009).	24
Figure 8. Examples of scenario terrains.	36
Figure 9. Example of virtual model positioning within a tetrad.	39
Figure 10. View of the VE display for experimental scenarios.	40
Figure 11. Example of a highlighted target.	43
Figure 12. Example view of the interface training scenario.	47
Figure 13. Mean detection accuracy across conditions per target behavior cue type: Standard error indicated by error bars.	53
Figure 14. Mean classification accuracy across conditions per target behavior cue type. Standard error indicated by error bars.	54
Figure 15. Median response time means for the clenched fists cue. Standard error indicated by error bars.	55

Figure 16. Median response time means for the slapping hands cue. Standard error indicated by error bars.	56
Figure 17. Median response time means across conditions per target behavior cue type. Standard error indicated by error bars.....	57
Figure 18. Percent change in detection accuracy per target behavior cue type across conditions. Standard error indicated by error bars.....	58
Figure 19. Percent change in classification accuracy per target behavior cue type across conditions. Standard error indicated by error bars.....	59
Figure 20. Percent change in median response time per target behavior cue type across conditions. Standard error indicated by error bars.....	60
Figure 21. Engagement Measure means between conditions for item seven, I like the challenge that using a virtual environment for behavior cue detection training provided. Standard error indicated by error bars.....	61
Figure 22. Immersion Measure means between conditions for item 15, The scenarios were challenging. Standard error indicated by error bars.....	62

LIST OF TABLES

Table 1. Target and non-target behavior cues for experimental stimuli.	38
Table 2. Number of target versus non-target events per scenario type per condition.	41
Table 3. Control group correlates between detection accuracy, engagement, and immersion.	63
Table 4. Control group correlates between classification accuracy, engagement, and immersion.	64
Table 5. Control group correlates between percent change performance metrics and immersion.	65
Table 6. Control group correlates between detection accuracy and presence.	66
Table 7. Control group correlates between classification accuracy and presence.	67
Table 8. Control group correlates between overall percent change in detection and classification accuracy and presence.	67
Table 9. Control group correlates between check six percent change metrics and presence.	68
Table 10. Control group correlates between wringing hands percent change metrics and presence.	69
Table 11. Massed Exposure correlates between percent change in detection accuracy and presence.	70
Table 12. Massed Exposure correlates between percent change in classification accuracy and presence.	71
Table 13. Highlighting group correlates between percent change performance metrics, engagement and immersion.	72

LIST OF ACRONYMS/ABBREVIATIONS

AAR	After Action Review
AOI	Area(s) of Interest
BEI	Biometrics-Enabled Intelligence
CCIR	Commander's Critical Information Requirements
COA	Courses of Action
CRT	Cathode Ray Tube
DOD	Department of Defense
HTA	Human Terrain Analysis
HVI	High-Value Individual
IKD	Intelligence Knowledge and Data
IO	Intelligence Operations
IP	Intelligence Personnel
ISR	Intelligence, Surveillance, Reconnaissance
IVD	Indirect Vision Display
I&W	Indications and Warnings
RAISR	Robot-Aided Intelligence, Surveillance, Reconnaissance
SBT	Scenario-Based Training
SBT-AID	Scenario-Based Training: Adaptive, Intelligent, Dynamic
UAS	Unmanned Aerial System
UCF	University of Central Florida
UGS	Unmanned Ground System
VE	Virtual Environment

CHAPTER ONE: INTRODUCTION

Since 2007, the U.S. Department of Defense (DOD) continues to revise and develop plans to increase the number of unmanned systems integrated into combat operations (DOD, 2013). This demand is spurred by desires to increase safety by removing Warfighters from direct line-of-fire and enhancing operational capabilities with technological support (DOD, 2013). Unmanned systems serve in several capacities including ordnance disposal, target acquisition, communications relay, persistent surveillance, air and ground reconnaissance, checkpoint assistance, and weapons delivery (Milburn, 2012; Army Research Laboratory, 2011; DOD, 2013; U.S. Army, 2008). Additionally, there is an emerging demand to enhance Intelligence, Surveillance, Reconnaissance (ISR) tasks with unmanned systems, herein referred to as Robot-Aided ISR (RAISR).

By 2024, the U.S. Army plans to modularize its force by reducing current Army divisions from 15,000 soldiers to brigades numbering 4000 soldiers, thus, increasing the overall number of units (U.S. Army, 2008). These changes in the organization of forces will necessitate the integration of systems that will support Joint ISR operations across more broadly dispersed units. The future modular force structure will require coordination and cooperation of ISR analysts across units to collectively analyze adversarial behavior to identify a baseline and implement preemptive, rather than reactive, actions in anticipation of attack (U.S. Army, 2008). A shift in modern warfare characteristics to irregular, urban combat situations presents specific requirements for emerging RAISR capabilities to enable detection of suspicious individuals from a safe position, distinguishing of behavioral characteristics indicating a suspicious individual's intent, and interpretation of a suspicious individual's behavior according to socio-cultural

characteristics (U.S. Army, 2008). Technological features to support these requirements may involve the employment of more sophisticated systems, cameras, and sensors to monitor the environment (DARPA, 2011; U.S. Army, 2008).

Along with the acquisition of new systems and capabilities, there emerges the added need to develop training requirements for instruction in system operation and specific task execution. There is limited documentation addressing specific training standards and objectives for ISR tasks utilizing unmanned systems (DOD, 2013). Much of the training doctrine that is well-defined addresses the air domain alone. The overarching focus of this research effort was to explore the utilization of unmanned systems in the ground domain to improve ISR quality and effectiveness for the identification of suspicious persons. Furthermore, the aim for this experiment was to investigate instructional solutions that promote effective and efficient RAISR task training.

The unmanned systems literature presents numerous concerns regarding the implications of incorporating unmanned systems on the evolution of ISR training. The U.S. Army's Concept Capability Plan for ISR (U.S. Army, 2008) and the Unmanned Systems Integrated Roadmap distributed by the DOD (DOD, 2013) describe specific questions addressing gaps in the unmanned system training research, which include:

1. What are the training implications of emerging unmanned system technologies to support ISR tasks?
2. What format of training will enhance development across the full spectrum of unmanned ISR capabilities?
3. Which type and what frequency of training will support skilled unmanned ISR operations?
4. What simulation-based resources are necessary for unmanned ISR training?

While the novelty of these emerging RAISR capabilities may necessitate some level of live training experience, simulation-based training platforms designed to represent real-world operational environments with a high degree of realism and meaningful practice have already been identified as a critical training requirement for unmanned ISR tasks (U.S. Army, 2008). One method to accomplish this is Scenario-Based Training (SBT) within a Virtual Environment (VE). To further enhance the effectiveness of training, instructional strategies may be integrated in SBT. This experiment applied a SBT framework to virtual training of a RAISR task and assessed the effectiveness of selected instructional strategies by evaluating measures of trainee performance and perception.

CHAPTER TWO: LITERATURE REVIEW

Intelligence, Surveillance, Reconnaissance

According to the U.S. Army definition, ISR is an enabling operation involving the planning, collection, analysis, and dissemination of data pertinent to the fulfillment of a Commander's Critical Information Requirements (CCIR) (U.S. Army, 2010). CCIR include any data needed to secure and retain situational understanding of an Area of Interest (AOI) and facilitate decision-making through informed selection of Courses of Action (COA) (U.S. Army, 2010). The U.S. Joint Forces definition expands the concept stating that ISR is a synergistic operation involving the coordination of available assets, sensors, and processing systems used to collect data (U.S. Army, 2008). ISR is considered a Combined Arms Operation utilizing an integrated network of assets from multiple echelons, both internal and external to the military, to actuate data collection efforts (U.S. Army, 2008; U.S. Army, 2009). As a collective term, ISR connotes an active process of gathering and interpreting information. However, to understand the full extent of ISR, it is befitting to discuss the distinct qualities of the individual terms (i.e., Intelligence, Surveillance, and Reconnaissance) comprising the acronym.

Intelligence

Intelligence is the more complex element of the ISR construct. Within the U.S. military context, Intelligence has a trifold definition, referring to Intelligence as a process, product, and personnel. As a *process*, Intelligence pertains to the actions involved to collect, integrate, analyze, and interpret data from the AOI, foreign nations, and hostile or potentially hostile forces (U.S. Army, 2010). The *product* of these processes is also called Intelligence and refers to the actual knowledge and data obtained (U.S. Army, 2010). Further, the term Intelligence is used to

designate the *personnel*, including analysts, Soldiers, and other assets, who execute ISR operations and manage data collected (U.S. Army, 2010). For the purposes herein, it is necessary to revise the labels of each definition in order to more readily identify intended meaning. Hereafter, Intelligence Operations (IO) will refer to the processes involved in the acquisition and management of data from the AOI. Intelligence Knowledge and Data (IKD) will refer to the knowledge and data products acquired as a result of IO. Finally, Intelligence Personnel (IP) will refer to individuals and organizations responsible for implementing IO and obtaining IKD.

Surveillance

The simple definition of Surveillance is the collection of data through observation. The specific U.S. military definition states that Surveillance entails employing visual, aural, and technological resources to systematically monitor locations, actors, and objects within designated space, surface, and subsurface areas (U.S. Army, 2007; U.S. Army, 2012b). The purpose of Surveillance is to support IO by maintaining persistent, uninterrupted observation of the designated AOI and by waiting for anomalies or changes in the environment to emerge (U.S. Army, 2012b). Surveillance aids the detection of critical changes in the state of the environment and status of entities of interest, thus enabling timely communication of Indications and Warnings (I&W) regarding adversarial actions, which may require shifts in tactical planning (U.S. Army, 2008; U.S. Army, 2012b). For the purposes of the present effort, the simplified definition of Surveillance as the collection of data through observation is sufficient.

Reconnaissance

Reconnaissance is similar to Surveillance in that it involves the collection of data through observation. However, unlike the passive and continuous observation of a designated AOI, which is characteristic of Surveillance operations, Reconnaissance operations involve active data collection from different locations within the AOI for a limited timeframe (U.S. Army, 2007; U.S. Army, 2012b). Additionally, while Surveillance operations often utilize a variety of visual, aural, and technological resources, Reconnaissance operations tend to heavily rely on human assets for gathering IKD (U.S. Army, 2007). However, with the increase in RAISR capabilities, the trend is shifting to incorporate more technological resources. IKD collected during Reconnaissance operations include information regarding the threat composition, strengths and weaknesses, capabilities, and resources of an adversary located beyond areas occupied by friendly forces (U.S. Army, 2008; U.S. Army, 2010). Additionally, CCIR may require collection of hydrographic, topographical, and meteorological data during Reconnaissance operations (U.S. Army, 2008).

ISR Perceptual Skills

ISR tasks are largely perceptual as they rely on the senses, primarily sight, to survey the environment. Visual perception involves the eyes' sensation of reflected light to detect objects and events in the environment (Bruce, Green, & Georgeson, 2003; Yantis, 2001). Within a military ISR context, visual perception is a critical ability for perceptual tasks involving threat detection, tactical decision making, and situation awareness (Carroll, Milham, & Champney, 2009). Perceptual skills are the acts and processes related to the performance of a perceptual task (Carroll, Milham, & Champney, 2009). Examples of perceptual skills applicable to ISR tasks

include attentional weighting, visual search, and pattern recognition (Carroll, Milham, & Champney, 2009; Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013).

Attentional weighting involves directing more attention (i.e., more weight) to cues that align with specific criteria and directing fewer visual resources to cues that do not fit specifications (Eckstein, Abbey, Pham, & Shimozaki, 2004; Rezec & Dobkins, 2004). Visual search involves visually scanning the environment in search of specific objects or features (Neisser, 1964). Pattern recognition is the interpretation and synthesis of a set of perceived characteristics or cues to identify familiar or discernible patterns or phenomena (Tarr, 2000). In a military ISR context, development of these perceptual skills is imperative for effective detection and interpretation of I&W that may signify the presence of a threat. U.S. Army IP use these perceptual skills either directly by dismounted infantry or indirectly via RAISR technologies.

Robot-Aided ISR

RAISR involves the utilization of remotely-operated, autonomous, or semi-autonomous unmanned systems to support ISR operations. A critical purpose for RAISR is to provide sustained and persistent Surveillance and Reconnaissance in order to maintain and increase battlespace awareness, which requires understanding the AOI status and its implications on Commanders' decision-making (DOD, 2013). The U.S. Army employs both Unmanned Aerial Systems (UASs) and Unmanned Ground Systems (UGSs) to conduct RAISR tasks. The current collection of U.S. Army unmanned systems with ISR support features spans a range of sizes, from compact, portable systems, such as the RQ-11B Raven UAS or Throwbot UGS which are rapidly deployable during short-term Reconnaissance missions, to much larger systems with

more advanced capabilities to support longer duration Surveillance operations, such as the RQ-7B Shadow UAS (Coba, 2010; Pearson, Moore, Ogdock, & Choi, 2013; Voth, 2004).

Despite the size and peripheral features of each system, a common feature is the use of Indirect Vision Displays (IVD) consisting of cameras and sensors mounted on the unmanned system that relay video surveillance footage to displays so that users may view the AOI and maintain situational understanding of the environment (Chen, et al., 2013). Current unmanned system IVD varieties include hand-held mobile devices, portable ruggedized laptops, and large Ground Control Stations housing multiple Cathode Ray Tube (CRT) or widescreen digital monitors (Coba, 2010; Pearson, Moore, Ogdock, & Choi, 2013; Voth, 2004).

An emerging feature undergoing research and development is Mind's Eye, an intelligent camera surveillance system that assists in the detection and interpretation of threats (MIT, 2010). Through the combination of computer vision and machine intelligence, the Mind's Eye system has the ability to detect objects and people in the environment, overlay visual aids such as a highlighted box to draw attention to cues, and apply meaningful text labels to describe what is "seen" (Bardu, et al., 2012; Bouma, et al., 2012; de Penning, den Hollander, Bouma, Burghouts, & d'Avila Garcez, 2012). One possible application of the Mind's Eye technology is to mount the cameras and sensors on unmanned systems to support decision-making during ISR missions (DARPA, 2011).

RAISR Perceptual Skills

Ultimately, unmanned systems and their support features are intended to improve the quality, effectiveness, and efficiency of ISR operations by enhancing the perceptual capacities of IP conducting such tasks. RAISR increases accessibility to unsafe or inaccessible AOI, thus,

broadening IP's perceptual scope and allowing them to get "eyes on the scene." The result is a greater quality and quantity of available IKD to support CCIR to obtain clear situational understanding.

UASs are able to provide a broader, overall view of the AOI and have been the front runner in Surveillance operations support for some time. However, due to an increasing shift toward irregular and urban combat environments, UGSs will likely be more valuable to fulfill emerging ISR capability requirements involving the detection and interpretation of suspicious human behavior and adversarial intent. In fact, supporting ground Surveillance and Reconnaissance with more UGSs has already been identified as a critical capability requirement for detecting High-Value Individuals (HVI) concealed within complex environments and urban populations (U.S. Army, 2008).

High-Value Individuals

HVIs are specific individuals whose presence, functions, and/or capabilities are critical to the orchestration of enemy operations or adversarial attacks (Fautua & Schatz, 2010; U.S. Army, 2012a). CCIR regarding HVIs requires analysis of the human terrain, which may include human behaviors, level of influence, location, physical description, function or rank within the enemy network, and degree of hostility (Fautua & Schatz, 2010; Schatz, Reitz, Nicholson, & Fautua, 2010; U.S. Army, 2012b). IP employ Human Terrain Analysis (HTA) techniques to assess physical geography and boundaries of HVI occupied areas, socio-cultural behavior, and environmental influences within the AOI to better understand the capabilities and motives that may impact HVIs' COA (U.S. Army, 2012a; U.S. Army, 2012b). In order to confirm a HVI's identity, IP may collect IKD of measurable physical features and behavioral characteristics, a

procedure referred to as Biometrics-Enabled Intelligence (BEI) (U.S. Army, 2010; U.S. Army, 2012a). Together, HTA and BEI contribute to a better understanding of the HVI's patterns of behavior which may contribute to I&W signaling the potential of a conflict or attack (Figure 1).

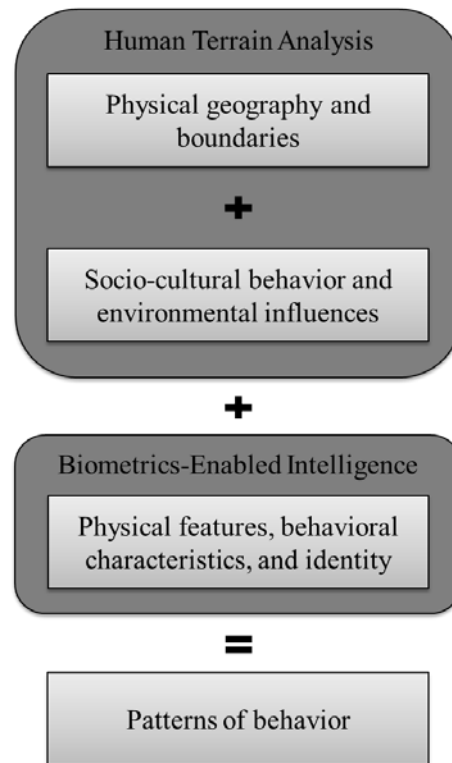


Figure 1. Types of Human Terrain Analysis and Biometrics-Enabled Intelligence data.

The U.S. Army identifies a need to advance RAISR technologies by incorporating HTA and BEI type capabilities (U.S. Army, 2008). However, these demands lack a clear theoretical foundation to delineate specific purposes and functions of HTA and BEI in ISR or RAISR operations. Interestingly, the elements of HTA and BEI closely align with another human terrain assessment strategy called Combat Profiling.

Combat Profiling

Combat Profiling is a strategy employed to analyze the human terrain through the appraisal of human behavior and interactions within the combat environment (Schatz, Reitz, Nicholson, & Fautua, 2010). This differs from the common notion of *profiling*, which is the appraisal of an individual's identity, intent, or potential involvement in adversarial operations based solely on observable features such as race, ethnicity, gender, physical or cultural characteristics, or clothing. Unlike the traditional concept of profiling, Combat Profiling is a more holistic approach to identifying HVIs within the AOI. Initially, an environmental and behavioral baseline is established by collecting IKD through Surveillance and Reconnaissance type operations (Ross & Militello, 2013). Next, environmental and behavioral changes are compared to the baseline and analyzed to determine whether or not the anomalies indicate the emergence of a potential threat (Salcedo, et al., 2013). The baseline and anomaly comparison procedure of Combat Profiling is a proactive approach to threat detection and conflict avoidance referred to as operating "left-of-bang" (Figure 2). Rather than reactive responses after the occurrence of an attack, Combat Profiling assists Commanders in taking preemptive COA to eliminate, capture, detain, contact, survey, or release a HVI before the catastrophic incident, or the "bang" (Ross & Militello, 2013; Flynn, 2010; Schatz, Reitz, Nicholson, & Fautua, 2010).

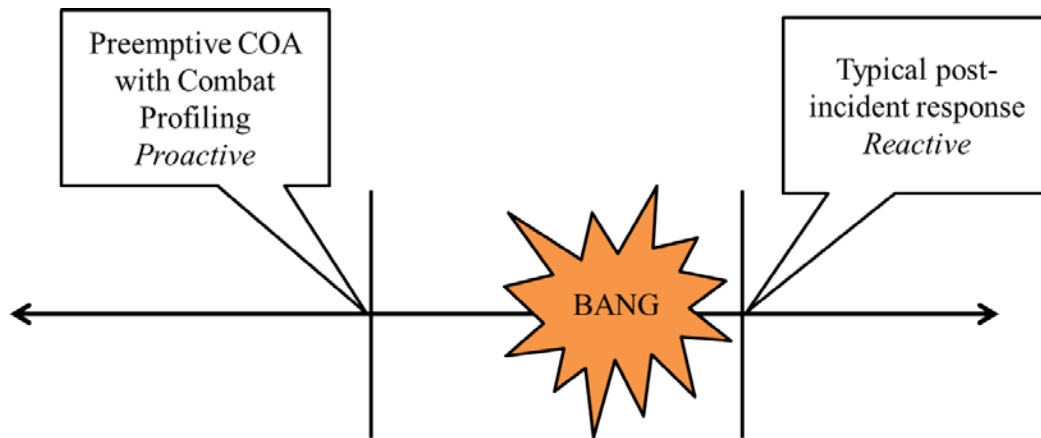


Figure 2. Proactive approach with Combat Profiling versus reactive response to an attack. Adapted from (Ross & Militello, 2013).

There are six distinct domains employed during Combat Profiling in order to holistically analyze the human terrain. These domains include: Geographics, Atmospheric, Proxemics, Biometrics, Kinesics, and Heuristics (Gideons, Padilla, & Lethin, 2008; Ross, Bencaz, & Militello, 2010). To provide a complete illustration of the Combat Profiling strategy, each domain is briefly described.

Geographics

The Geographics domain involves the survey of an environment's physical geography and the analysis of human interaction with that environment (Gideons, Padilla, & Lethin, 2008; Ross, Bencaz, & Militello, 2010). For example, physical features of the environment may allow for free access to all humans in some areas, while access to other areas is restricted by the physical terrain, such as mountain ranges, or by socio-cultural boundaries imposed by an organized group, such as insurgents who secure an area by force for tactical advantage.

Atmospherics

Atmospherics refers to the perceived mood or ambience resultant of environmental features and characteristics (Gideons, Padilla, & Lethin, 2008; Ross, Bencaz, & Militello, 2010). In general, atmospheric cues are detected by the senses including: sight, sound, smell, taste, and touch. Initially, atmospheric data contribute to the establishment of an environmental baseline. Afterward, atmospheric changes may signal the presence of an environmental anomaly, such as smelling a strong odor on a street where there previously was no such odor.

Proxemics

The Proxemics domain addresses the spatial relationships between the people within an environment (Gideons, Padilla, & Lethin, 2008; Ross, Bencaz, & Militello, 2010). Proxemics is the study of people within groups or networks. It focuses on the separation or closeness maintained by individuals and how that relationship influences socio-cultural interactions. For example, individuals may exhibit subservience by increasing the distance in the presence of a dominant or revered person in the community.

Biometrics

The Biometrics domain involves the analysis of involuntary physiological responses of the human body and the interpretation of those responses to identify an individual's affective state, such as anger or shame, or other physical influence, such as drugs or alcohol (Gideons, Padilla, & Lethin, 2008; Ross, Bencaz, & Militello, 2010; Salcedo, et al., 2013). This differs from measurable physical features and behavioral characteristics of BEI data from ISR, which are used to verify a target's identity. Examples of biometric cues in Combat Profiling include:

sweating and increased heart rate when nervous, protruding veins in the face or neck and facial flushing when angry, dilated pupils when under the influence of drugs, blushing when embarrassed, and facial pallor when frightened or ill. It is critical that context is carefully considered when analyzing biometric cues as environmental factors, such as the weather or climate, or other physiological factors, such as medical conditions, may influence the portrayal of these cues.

Kinesics

Kinesics comprises the study of nonverbal body cues and actions that convey meaning (Birdwhistell, 1970). Within Combat Profiling, the Kinesics domain involves the analysis of body language, facial expressions, and gestures (Gideons, Padilla, & Lethin, 2008; Ortiz, Maraj, Salcedo, Lackey, & Hudson, 2013; Ross, Bencaz, & Militello, 2010). Analyzing kinesic cues assists in identifying an individual's affective state, such as clenched fists signifying anger or covering the mouth with one hand when lying (Ortiz, Maraj, Salcedo, Lackey, & Hudson, 2013).

Heuristics

Heuristics refers to the brain's tendency to create generalizations and generate rules based on perceived environmental patterns (Gideons, Padilla, & Lethin, 2008; Ross, Bencaz, & Militello, 2010). Using Heuristics as a frame of reference can promote rapid decision-making. For example, the heuristic domain is beneficial during initial baseline formation. Through the combination of the other five domains of Combat Profiling, recognizable patterns begin to emerge, which assist in the generalization of expected patterns of life (Figure 3). However, IP must be vigilant because this cognitive phenomenon may also cause oversight of critical threat

cues, such as assuming an individual seen standing in the same place every day is not involved in the enemy network.

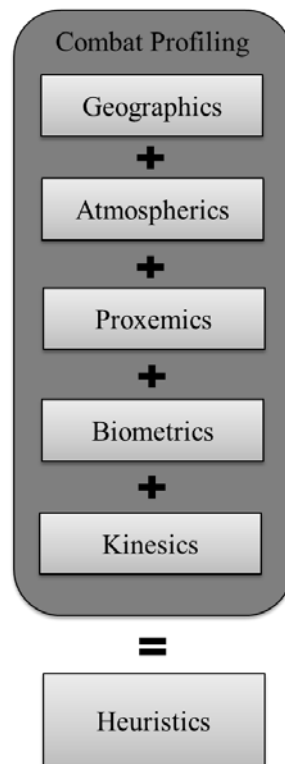


Figure 3. The six domains of Combat Profiling.

Role of Combat Profiling in ISR

The Combat Profiling domains are considered culturally agnostic, meaning the types of data cues observed are applicable in any cultural or warfare setting from current rural Middle Eastern conflicts to anticipated conflicts in more urbanized, irregular environments (Lackey & Salcedo, 2014; Spiker, Williams, Johnston, & Lethin, 2010). Incidentally, cue data observed and collected via Combat Profiling, regardless of culture or terrain, closely aligns with the categories of HTA and BEI data for ISR operations (Figure 4).

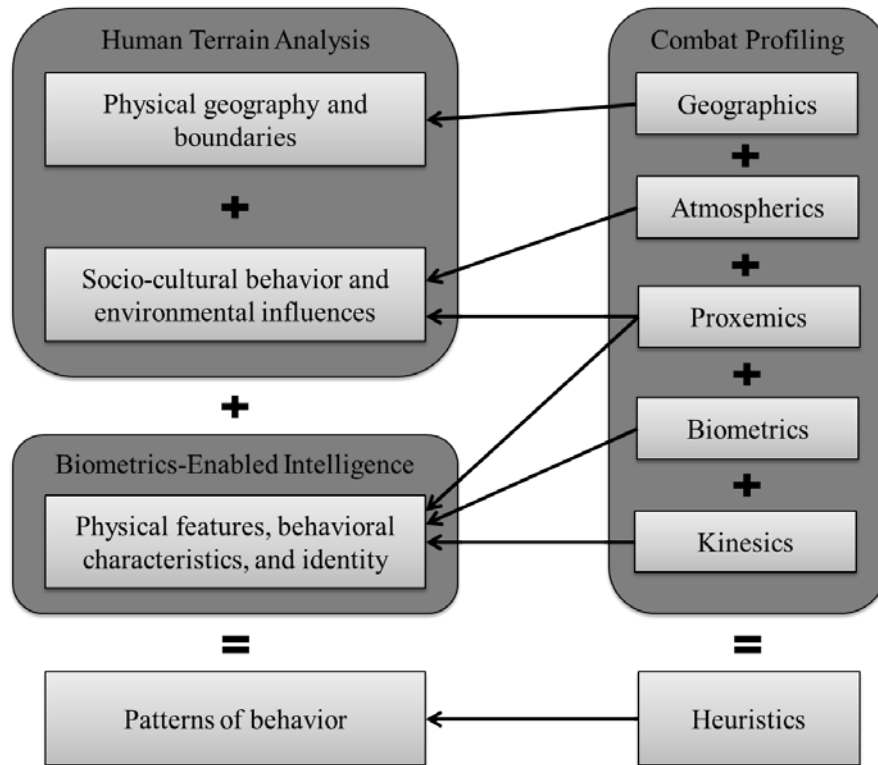


Figure 4. Alignment of ISR data types for HVIs and the six Combat Profiling domains.

Geographic cues contribute to the identification of the physical geography and boundaries of areas inhabited by HVIs. Atmospheric and proxemic cues are indicators of socio-cultural behavior and environmental influences. Physical features, behavioral characteristics such as affective state, and HVI identity may be determined by employing the Biometrics, Kinesics, and Proxemics domains. Finally, by combining all six domains, patterns of behavior emerge, thus, promoting the generation of Heuristics to support the Commander's selection of COA.

Extensive Combat Profiling research and development efforts have primarily focused on the application of the strategy from the perspective of a remote observation post (Colombo, Dolletski-Lazar, Coxe, & Tarr, 2012; Schatz, Folsom-Kovarik, Bartlett, Wray, & Solina, 2012; Schatz, Reitz, Nicholson, & Fautua, 2010). However, there is no evidence suggesting that

Combat Profiling is not applicable to other observational perspectives, such as dismounted infantry or IVDs. Combat Profiling even utilizes the same perceptual skills required for ISR and RAISR tasks as the tools to detect cues and threats in the human terrain. Clearly, the similarities support leveraging Combat Profiling as a strategy to define HTA and BEI criteria when collecting HVI data during ISR or RAISR tasks (Figure 5).



Figure 5. ISR task, strategy, tool hierarchy.

More recent empirical investigations have applied the Kinesics domain of Combat Profiling to define the criteria for identifying potential HVIs during a RAISR behavior cue analysis task, which employs attentional weighting, visual search, and pattern recognition skills to detect specific behavior cues (e.g., kinesic cues) and classify those cues to determine a HVI's affective state (e.g., aggressive, nervous, fear, anxiety, etc.) (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013; Lackey & Salcedo, 2014; Salcedo, Lackey, & Reinerman-Jones, 2014). Select kinesic cues representing aggressive or nervous affects were depicted virtually using animated virtual character models in a simulation-based platform for the purpose of training the behavior

cue analysis task. Behavioral indicators of aggressiveness included clenching the fists and slapping the hands together. Clenched fists are evident by the curling and squeezing of the fingers into the palms of the hands signifying a stress response to feelings of anger (Givens, 2002). Slapping the back of one hand into the palm of the other hand is a sign of emphasis that may be evidence of agitation (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013). The selected nervousness cues included wringing the hands together and check six behavior. Wringing the hands involves repeatedly alternating the clasping and squeezing of one hand and fingers with the other hand (Navarro & Karlins, 2008). “Check your six” originated as a warning to fighter pilots when enemy aircraft approached them from behind (Dalzell, 2009). The phrase has evolved as a slang reference for looking behind oneself (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013). People may look over their shoulders and turn around if they feel nervous or anxious that someone may be watching them or their actions (Coover, 1913).

Identifying aggressive and nervous behavior cues may indicate the presence of suspicious persons with malicious intent and signify the threat of an impending attack from adversarial forces. Individuals may exhibit aggressive cues when encountering a situation or other individual that is displeasing, such as a civilian angered by the presence of adversarial forces in the area. Nervous cues may indicate individuals who are in distress due to hostile behavior nearby or even those attempting to conceal deceptive actions, such as an adversary worried about being revealed prior to an attack. The ability to accurately detect and classify aggressive and nervous behavior cues may provide valuable insight concerning a HVI’s influence within an AOI.

These prior experiments assessed behavior cue analysis ability when presented with a series of training events consisting of individual targets exhibiting the behavior cues (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013; Lackey & Salcedo, 2014; Salcedo, Lackey, &

Reinerman-Jones, 2014). Individual targets were detected at 94% accuracy and classified at 85% accuracy with an average response time of 3.56 seconds. While this method was necessary in order to establish an empirical baseline for the behavior cue analysis task and the selected kinesic cues, it is unrealistic that IP will monitor and assess one individual at a time during real-world RAISR operations. Therefore, a logical progression is to provide a more realistic level of complexity by increasing the number of entities per training event, thus, creating groups of potential targets. Past research suggests that competitive behaviors more often arise in groups of at least four, or tetrads, opposed to smaller groups (Benenson, Nicholson, Waite, Roy, & Simpson, 2001). Further, individuals participating in competitive activities have a greater potential of demonstrating hostile or aggressive behavior in larger groups, such as tetrads and six-person groups, opposed to smaller groups, such as dyads (Eastin, 2007). Therefore, within the context of behavior cue analysis, aggression may be more relevant during the observation of tetrads opposed to smaller groups or individuals. Additionally, tetrads are applicable for the presence of nervous behavior cues, which are a realistic counter response to aggressive behavior. Aggressive or hostile behaviors are often met by submissive, nervous, or evasive responses from subordinates, targets of the adverse behavior, or individuals located near the aggressor(s) (Orford, 1986; Potegal & Knutson, 1994).

In RAISR operations, picking out specific behavior cues from a group via an IVD will require IP to rely even more on their perceptual skills. Realistic and focused training has shown to improve perceptual skill performance during task execution (Hale, et al., 2012; Seitz & Dinse, 2007). Therefore, in order to provide more realistic skill application opportunities, IP utilizing RAISR technologies would benefit from training that provides the same perspective as the IVD of the system. The current and emerging varieties of IVDs transition well into simulation-based

training utilizing digital displays such as mobile devices, laptops, and desktop PCs (Salcedo, Lackey, & Maraj, 2014). However, simulation-based training environments alone will not fulfill the complex training needs of RAISR tasks.

Training

Whether the simulated environment is live, virtual, or constructive, the term *simulation-based training* simply denotes the utilization of a simulated environment for instruction and skill practice, not a specific instructional approach (Martin, Hughes, Schatz, & Nicholson, 2010). Past research found that the use of simulations without appropriate instructional strategies often results in negative training due to ineffective and inefficient simulation training quality (Oser, Gualtieri, Cannon-Bowers, & Salas, 1999). The shortcomings of early simulation-based training attempts spurred research addressing the structure of content within simulation exercises (Martin, et al., 2009). While traditional classroom instruction consists of a series of lessons, it was found that effective simulation-based training should consist of a series of related training exercises in which trainees apply their acquired knowledge and skills (Oser, Cannon-Bowers, Salas, & Dwyer, 1999; Lyons, Schmorrow, Cohn, & Lackey, 2002). This led to the foundation of the Scenario-Based Training (SBT) approach to simulated training.

Scenario-Based Training

SBT emerged as a design approach for simulations involving the purposeful instantiation of learning and practice opportunities that elicit desired psychological states during scenario events (Martin, et al., 2009; Martin, Hughes, Schatz, & Nicholson, 2010; Martin, Schatz, Hughes, & Nicholson, 2010). Scenario content is planned based on skill inventories and archived

performance data or desired performance outcomes, which are used to develop a task list. The task list informs the composition of learning objectives and critical competencies, which are in turn utilized to derive scenario events and scripts. Scenario events are developed through the combination of pre-defined triggers and adaptations to generate scenario vignettes, which are sequenced together to create a training scenario. Then, performance measures and standards of skill mastery, linked directly to the objectives and events, are developed and implemented during the training exercise. Performance is diagnosed during scenario execution and subsequently analyzed to identify trainee strengths and weaknesses. Results of performance analysis are used to formulate assessment feedback for After Action Review (AAR) and archived for reference during future training instances. A benefit of SBT is that all stages of the training process from planning to execution to assessment are closely linked so that the objectives, content, and desired performance outcomes are consistent (Figure 6).

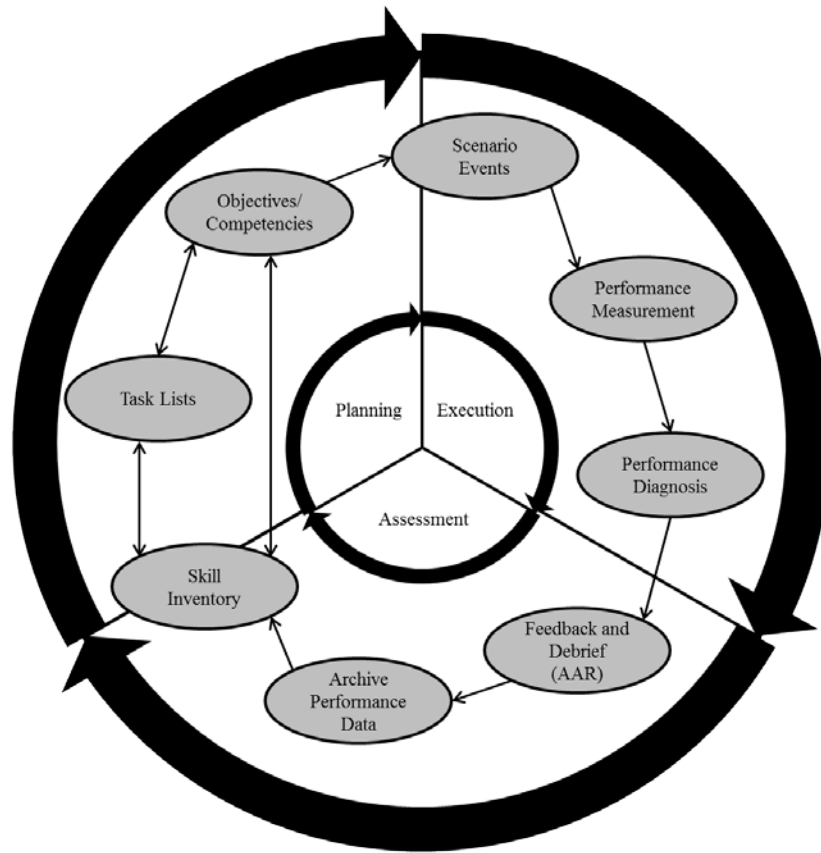


Figure 6. Components of the Scenario-Based Training framework. Adapted from (Fowlkes, Dwyer, Oser, & Salas, 1998; Lyons, Schmorow, Cohn, & Lackey, 2002).

Previous research indicates that the SBT approach increases training effectiveness when applied to the development of simulation content for training procedural tasks (Dunne, Schatz, Fiore, Martin, & Nicholson, 2010). Since training objectives are directly linked to specific scenario events and desired performance outcomes, the systematic and prescriptive design approach of SBT works well in domains with predictable, repeatable events. However, the dynamics of modern warfare have increasingly shown that “unpredictability is the most predictable characteristic of military operations” (Fletcher, 2004). This unpredictability has necessitated an expansion and diversification of training requirements beyond procedural tasks to

include individualized strategies for higher-order thinking and perceptual skill development (Carroll, Milham, & Champney, 2009; Dunne, Schatz, Fiore, Martin, & Nicholson, 2010). Formerly, such skills were acquired over time on a domain specific basis through the apprenticeship of junior officers with knowledgeable senior ranking mentors (Becker & Schatz, 2010). While it is effective, apprenticeship is inefficient considering the rapid evolution and escalation of the modern warfare climate (Becker & Schatz, 2010; Nicholson, Schatz, Stanney, & Lackey, 2009). Therefore, to account for the intricacies of emerging demands and training requirements, the SBT framework received further refinement through the development of the Scenario-Based Training: Adaptive, Intelligent, Dynamic (SBT-AID) approach.

Scenario-Based Training: Adaptive, Intelligent, Dynamic

SBT-AID extends the SBT model to include “intelligent tutoring components, scenario-based instructional simulations, dynamic scenario generation capabilities, content authoring support, and an integrated pedagogical framework” (Nicholson, Schatz, Stanney, & Lackey, 2009). The SBT-AID approach divides the training timeline into pre-, during-, and post-training activities (Figure 7). *Pre-training* includes completion of a task analysis, access of stored trainee profiles, selection of training objectives and instructional strategies, delivery of initial instruction, selection of the simulated environment, and generation of training scenarios. The *during-training* phase is where execution of the scenario occurs, which involves in-simulation performance assessment and diagnosis, real-time adaptation of the scenario to meet changing trainee needs, and presentation of instructional assistance to guide learning. After the trainees complete the scenario, the *post-training* activities commence with the diagnosis of overall

performance, instructional feedback and review of scenario performance, and documentation of performance and trainee profile data for application and reference in future exercises.

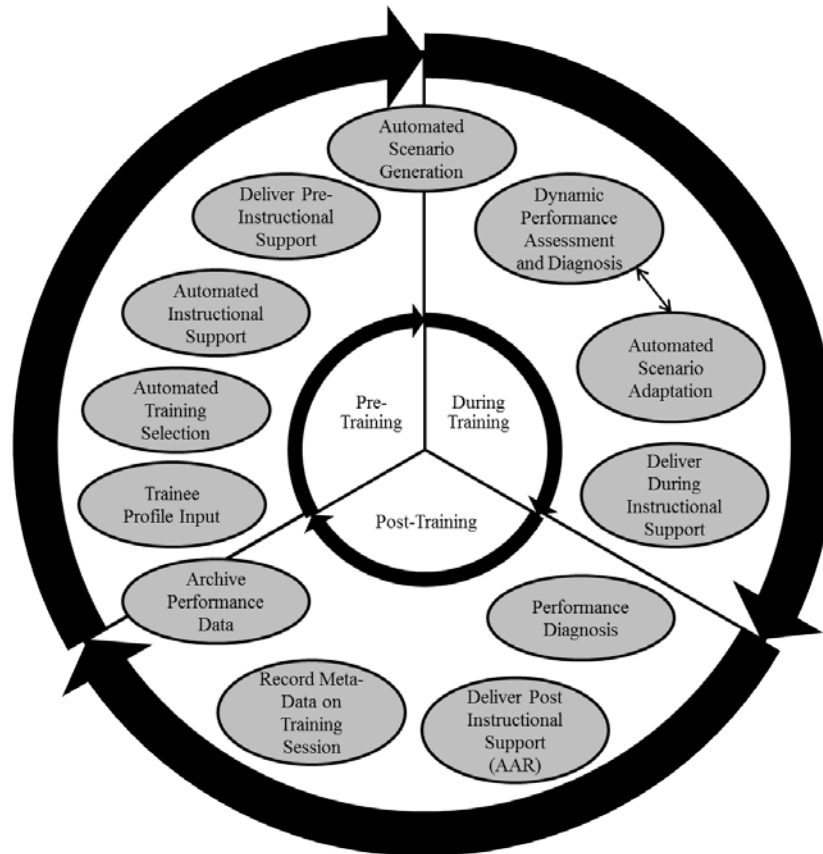


Figure 7. Components of Scenario-Based Training: Adaptive, Intelligent, Dynamic. Adapted from (Nicholson, Schatz, Stanney, & Lackey, 2009).

SBT-AID follows a systematic structure similar to SBT with the added attempt to embed the process with instructional supports that foster more rapid acquisition and improvement of higher-order thinking and perceptual skills formerly developed through apprenticeship (Nicholson & Schatz, 2010; Nicholson, Schatz, Stanney, & Lackey, 2009). SBT-AID introduces the importance of incorporating instructional strategies to enhance the effectiveness of the SBT framework when applied to training within simulated environments. Therefore, instructional

design research should respond with empirical evidence to support the selection and implementation of targeted instructional strategies to address specific skill needs.

Instructional Strategies

There are numerous instructional strategies described in the training and education literature from various domains; however, there are a limited number of empirical investigations assessing the integration of selected instructional strategies within a SBT setting. SBT has primarily been implemented for the training of procedural skills (Dunne, Schatz, Fiore, Martin, & Nicholson, 2010). More recent evidence supports the application of SBT for training higher-order cognitive skills, such as decision-making, planning, problem solving, and metacognition, with the Metacognitive Prompting and Contrasting Cases instructional strategies (Nicholson, Fiore, Vogel-Walcutt, & Schatz, 2009; Nicholson & Schatz, 2010; Becker & Schatz, 2010; Dunne, Schatz, Fiore, Nicholson, & Fowlkes, 2010; Fiorella, Vogel-Walcutt, & Fiore, 2012).

Metacognitive Prompting involves questioning or cueing learners to reflect on or self-monitor their own thought processes or metacognition (Fiorella, Vogel-Walcutt, & Fiore, 2012). The Metacognitive Prompting instructional strategy has been shown to improve Call for Fire decision-making performance in a U.S. Marine Corps Fire Support Team (Fiorella, Vogel-Walcutt, & Fiore, 2012). In medical training research, Metacognitive Prompting has been shown to improve visual search ability (Nodine, Mello-Thomas, Kundel, & Weinstein, 2002).

Contrasting Cases involves comparing and contrasting a set of related examples, or cases (Fowlkes, Norman, Schatz, & Stagl, 2009). The Contrasting Cases strategy has also been assessed for effectiveness in Call for Fire training resulting in recommendations to incorporate the strategy to support procedural skills associated with the task (Vogel-Walcutt, Marshall,

Schatz, Dolletski-Lazar, & Nicholson, 2011). Analyzing related cases may also support the development of pattern recognition skills because it requires the identification of commonalities, or patterns, across examples.

Regarding SBT for perceptual skills, Carroll, Milham, and Champney (2009) conducted an in-depth review of instructional strategies applied to perceptual skills training in the military domain and identified several strategies applicable to the development of attentional weighting, visual search, and pattern recognition including: Scaffolding, Massed Exposure, Minimum Stimulus, and Highlighting.

Scaffolding involves the decomposition of a task into its basic steps and the gradual release of task responsibility to the learner (Carroll, Milham, & Champney, 2009; Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013; van de Pol, Volman, & Beishuizen, 2010). In other words, the Scaffolding strategy helps learners construct knowledge by mastering basic skills first, and then gradually adding more complex skills until the task can be executed independently. The gradual progression from skill to skill may promote the recognition of familiar patterns (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013). Scaffolding has been linked to Lev Vygotsky's concept of the Zone of Proximal Development, which is the difference between a learner's actual performance and his/her potential performance achieved via instructor guidance (Sanders & Welk, 2005; van de Pol, Volman, & Beishuizen, 2010; Vygotsky, 1978). In the military domain, instructor modeling and guidance of visual search skills assists in scaffolding learners to apply expert techniques for visual search (Carroll, Milham, & Champney, 2009).

Massed Exposure refers to exposure to a high volume of stimuli or training events concentrated within one or a few exposure sessions (Carroll, Milham, & Champney, 2009; Hirumi & Stapleton, 2009). Military training research has identified Massed Exposure as a

perceptual skills training strategy to improve attentional weighting for threat cues by engaging soldiers in a variety of practice environments containing high volumes of threat events presented within a condensed timeframe (Carroll, Milham, & Champney, 2009). Empirical investigations have primarily applied the strategy to the training of motor skills with a majority of the studies reporting that Massed Exposure, also called Massed Practice, is ineffective at improving motor skill development (Donovan & Radosevich, 1999). However, some theories suggest that Massed Exposure may have a more positive impact on cognitive and perceptual skill development (Donovan & Radosevich, 1999). There is also evidence from sports psychology research to support the use of Massed Exposure to improve pattern recognition skills (Williams & Ward, 2003).

Minimum Stimulus is somewhat of an antithesis to the Massed Exposure strategy. As the inverse to Massed Exposure, Minimum Stimulus involves the presentation of a reduced number of target stimuli and/or non-target distractors (Fiore, Scielzo, Jentsch, & Howard, 2006; Kass, Herschler, & Companion, 1991). In a battlespace awareness context, saliency of cues intended to promote development of pattern recognition has shown to be greater when the Minimum Stimulus strategy is applied in a simulation-based training environment (Carroll, Milham, & Champney, 2009; Kass, Herschler, & Companion, 1991). The Minimum Stimulus strategy may also present a more realistic level of difficulty by reducing the volume of target stimuli to be consistent with actual or estimated target probabilities in the real-world environment (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013).

Highlighting, also called Exogenous Orienting, is used to orient trainees' visual resources to identify the emergence of a cue, object, image, or other event significant to the training context (Carroll, Milham, & Champney, 2009). Sports psychology often applies Highlighting as

an attentional weighting method to improve reaction time or anticipatory behaviors by directing players' attention to specific cues, events, or actions that arise during game play (Fuchs, McNevin, Ritter, Toole, & Wulf, 2000; Hagemann, Strauss, & Canal-Bruland, 2006; Jackson & Farrow, 2005; Poulter, Jackson, & Berry, 2005). Highlighting cues with the addition of a non-content feature, such as a spotlight or circle on the target object, has been shown to improve visual search skills in driver training (Chapman, Underwood, & Roberts, 2002; Underwood, 2007) and human anatomy learning (de Koning, Tabbers, Rikers, & Paas, 2010; de Koning, Tabbers, Rikers, & Paas, 2007).

Training Effectiveness Evaluation

In order to identify which SBT instructional strategy is optimal for training a specific task domain or task type, training effectiveness should be assessed along relevant and measurable factors. The enduring standard for training effectiveness evaluations across multiple training domains is Kirkpatrick's four level evaluation framework (Salas & Cannon-Bowers, 2001; Arthur, Jr., Bennett, Jr., Edens, & Bell, 2003; Kotnour, Landaeta, & Lackey, 2013). Kirkpatrick's model identifies learning, reaction, behavior, and results as the critical factors, or levels, for evaluating the effectiveness of training programs (Kirkpatrick, 1959; Kirkpatrick, 1976; Kirkpatrick & Kirkpatrick, 2006). The learning level assesses trainees' skill and knowledge acquisition by conducting performance tests (Kirkpatrick & Kirkpatrick, 2006; Arthur, Jr., Bennett, Jr., Edens, & Bell, 2003). The reaction level includes the trainees' perception and level of satisfaction with the training experience (Kirkpatrick & Kirkpatrick, 2006; Swanson & Sleezer, 1987). Evaluating the behavior level involves assessing trainees' application of the new knowledge and skills acquired during training to the real-world setting

(Kirkpatrick & Kirkpatrick, 2006; Schumann, Anderson, Scott, & Lawton, 2001). The final level evaluated is the results level, which assesses the impact of the training experience on improving performance quality of the organization, usually over a longer period of time (Arthur, Jr., Bennett, Jr., Edens, & Bell, 2003; Schumann, Anderson, Scott, & Lawton, 2001). Although assessment across all four levels provides a thorough evaluation of training effectiveness for fully developed training systems, an initial focus on only the reaction and learning factors during exploratory instructional design research may provide sufficient data to generate recommendations for conceptual, undeveloped, or future training systems.

Performance

One intention of training is to improve a trainee's ability to conduct a specific task, an obvious factor to assess is the impact of the training on performance, or the learning level of training effectiveness evaluation. For a RAISR task, such as behavior cue analysis, critical performance criteria involve the level of accuracy and speed at which targets are detected and classified (Lackey & Salcedo, 2014; Salcedo, Lackey, & Reinerman-Jones, 2014). In an actual combat environment, the accuracy and speed of identifying HVIs affect the success and safety of the mission (Lackey & Salcedo, 2014). Accurate and early analysis of HVI behavior aligns with the emerging ISR capability requirement to identify threats from a safe distance and to distinguish and interpret behavioral and socio-cultural indicators of intent (U.S. Army, 2008). Accuracy may be measured objectively by calculating the percentage of correct responses during performance testing. Speed of performance may be measured by recording trainees' response time, or the amount of time it takes to correctly respond to test items or events.

When conducting evaluations at the learning level, it may be necessary to reduce the risk of confounds due to the performance test design. For example, difficulty levels may affect trainee workload, which in turn may impact performance and skew evaluation results. Easy performance tests may not be sensitive enough to adequately assess the impact of the training on learning, while excessively challenging tests may risk the assumption that the training was ineffective. Therefore, a moderate level of difficulty may be appropriate for performance tests. In SBT, increasing or decreasing the number of training events often alters the level of difficulty (Martin, Schatz, Hughes, & Nicholson, 2010). Recent evidence from workload research found that a scenario event rate of 30 events per minute may induce a moderate level of workload during a threat detection task (Abich, Taylor, & Reinerman-Jones, 2013).

Engagement, Immersion, and Presence

At the reaction level, evaluation criteria must address trainee perception or satisfaction factors that are relevant to the task domain and desired training experience (Kirkpatrick & Kirkpatrick, 2006). In SBT platforms, relevant constructs for trainee reaction evaluation include engagement, immersion, and presence.

Engagement is defined as the level of involvement pursuant of a learner's degree of interest in the task, content, or media (Charlton & Danforth, 2005). Findings from training and education research suggest that learner engagement is positively correlated to learner satisfaction (Van den Berg, et al., 2007; Scott, 2008; Wefald & Downey, 2009; Lin, 2009; Havice, Davis, Foxx, & Havice, 2010; Levett-Jones, et al., 2011). Therefore, higher levels of engagement may be an indicator of greater satisfaction with the training.

In virtual simulation research, immersion often refers to the level of sensory fidelity provided by the simulation technologies (e.g., displays, rendering software, etc.) (Slater, Linakis, Usoh, Kooper, & Street, 1996; Bowman & McMahan, 2007). Measures of immersion characterize the construct as the sensation of “losing oneself” in an experience due to highly focused attention on the task (Jennett, et al., 2008; Witmer & Singer, 1998). The assumption is that this cognitive absorption in the task is induced by the technology (Jennett, et al., 2008; Slater, Linakis, Usoh, Kooper, & Street, 1996; McMahan, 2003). Simulation and game-based training research indicate that immersion and engagement are highly related constructs and together may impact satisfaction (Douglas & Hargadon, 2001; Murphy, 2011).

Presence, in simulation-based environments, is described as the sense of experiencing the simulated environment versus the physical environment (Witmer & Singer, 1998). In other words, presence is a conscious state of feeling as if one is a part of, or present in, the simulated environment to the exclusion of the real-world environment (Slater, Linakis, Usoh, Kooper, & Street, 1996; McMahan, 2003). Results from distributed learning and virtual world research suggest that a greater sense of presence increases learner satisfaction with the experience (Zhang & Zigers, 2009; Bulu, 2012).

Common methods to assess perceptions and satisfaction include self-reporting measures, such as surveys and questionnaires, to collect trainees’ personal evaluation of the experience (Arthur, Jr., Bennett, Jr., Edens, & Bell, 2003). Previous studies assessing SBT for behavior cue analysis of individual targets utilized modified versions of the engagement scale by Charlton and Danforth (2005), the immersion measures by Jennett et al. (2008), and the presence questionnaire by Witmer and Singer (1998) to evaluate perceptions of the training experience (Ortiz, Maraj, Salcedo, Lackey, & Hudson, 2013; Salcedo, Lackey, & Maraj, 2014; Lackey, Maraj, & Barber,

2014). Engagement, immersion, and presence have shown to impact performance (Csikszentmihalyi, 1990; McNamara, Jackson, & Graesser, 2009; Jennett, et al., 2008; Witmer & Singer, 1998). Therefore, SBT instructional strategies that promote higher levels of engagement, immersion, and presence may correlate to better performance outcomes.

Purpose of Present Experiment

The purpose for this research effort is to investigate the use of a UGS to assist in ISR tasks. As research and development regarding RAISR capabilities continues to expand, the demand for training and education solutions will emerge. Therefore, the goal for this study was to expand the SBT and instructional design literature related to the training of perceptual skills involved in RAISR operations. The specific RAISR task was the identification of HVIs within a group through behavior cue analysis. Furthermore, due to a lack of substantial empirical evidence evaluating the effectiveness of specific instructional strategies for SBT, this experiment assessed two instructional strategies, Massed Exposure and Highlighting, identified to improve attentional weighting, visual search, and pattern recognition skills. Massed Exposure was selected due to its expected capacity to improve attentional weighting and pattern recognition of specific cues. There is also a substantial gap in the instructional strategy literature investigating the application of Massed Exposure for perceptual skills training that this experiment addressed. Highlighting was selected for its ability to improve visual search skills when identifying a specific cue, object, or signal from a group or cluster of items or information. Also, the use of non-content features in Highlighting to orient learners' attention during a practice scenario may be designed for consistency with the indication features applied in intelligent surveillance systems, such as Mind's Eye. Therefore, the specific objective of the present experiment was to

empirically assess the effectiveness of the Massed Exposure and Highlighting strategies applied to SBT for behavior cue analysis training. Effectiveness was evaluated by analyzing the impact of the instructional strategies on post-test performance outcomes, including detection accuracy, classification accuracy, and response time, and on perceptions of the level of engagement, immersion, and presence during training exercises. Due to a lack of sufficient theoretical and empirical evidence to support any significant advantage in effectiveness of one strategy over the other, the following null hypotheses (H_0) were tested:

- There is not a significant difference between the Massed Exposure and Highlighting instructional strategies versus a Control for post-test performance outcomes.
- There is not a significant difference between the Massed Exposure and Highlighting instructional strategies versus a Control for training perceptions.
- There are no significant relationships between post-test performance outcomes and training perceptions for the Massed Exposure and Highlighting instructional strategies versus a Control.

CHAPTER THREE: METHODOLOGY

Participants

Participation was restricted to individuals between 18 and 40 years of age with U.S. Citizenship status. Since the task involved visual discrimination, normal vision or corrected to normal vision when wearing prescription lenses was required. Additionally, to maintain consistency with prior experimentation and current U.S. Army vision requirements, participation required full color vision as indicated by the Ishihara's Tests for Colour Deficiency (Ishihara, 2013). Finally, individuals who participated in the preceding experiment (Lackey & Salcedo, 2014; Salcedo, Lackey, & Maraj, 2014) were restricted from volunteering due to the similarity of the training content and experimental task.

Total participants included 123 volunteers from the University of Central Florida (UCF) campus and affiliated organizations. The data of five participants were excluded due to technical issues, failure to meet the proficiency requirement of the experimental task, and electing to discontinue participation. The remaining sample of 118 included 58 females and 60 males, ages 18 to 33 years ($M = 22.13$, $SD = 3.15$). Participants were randomly assigned to one of three group conditions: Control ($n = 39$), Massed Exposure ($n = 39$), and Highlighting ($n = 40$).

After fulfilling institutional consent requirements, participants were asked a series of pre-experiment questions to record use of alcohol, sedatives, anti-psychotic drugs, or anti-depressants within 24 hours or caffeinated substances within two hours prior to the experiment start time. Responses to these questions did not contribute to exclusion criteria, but were documented for reference during performance data analyses to account for potential outliers, if needed.

Compensation included a choice between monetary payment at a rate of ten U.S. dollars per hour or class credit in approved UCF courses awarded at the discretion of the associated course professor. The experiment duration was approximately four and one-half hours.

Experimental Testbed

A series of VE scenarios were developed using the Virtual Battlespace 2 (VBS2) version 2.0 software. VBS2 is a VE development software program currently used by the U.S. Army for simulation-based training and mission rehearsal (Ortiz, Maraj, Salcedo, Lackey, & Hudson, 2013; Bohemia Interactive, 2014). VBS2 2.0 provides a 3D virtual model library of objects, buildings, vehicles, animations, and geotypical terrains for customizable scenario development.

The VE scenarios were designed to simulate camera surveillance footage from an autonomous UGS collected during an ISR mission to detect and classify human behavior cues. Scenarios were displayed on a 22 inch (16:10 aspect ratio) computer monitor. The VBS2 2.0 customizable camera functionality allowed for the designation of a specified camera height and movement along predetermined waypoints in a specified direction and speed. For consistency with current and emerging robotics technologies, the virtual UGS camera height was set to one meter above the ground surface and traveled forward at one and a half meters per second (Mykoniatis, Angelopoulou, Soyler, Kincaid, & Hancock, 2012; MARCbot, 2010).

Scenario terrains included two settings: Middle Eastern and Culturally Agnostic (Figure 8). Culturally Agnostic refers to a non-geotypical setting that is representative of environments that may be found in many urbanized areas around the world. Building and virtual character models placed in the Middle Eastern setting are consistent with those found in real-world counterparts. Additionally, building and virtual character models placed in the Culturally

Agnostic settings are representative of what may be found in many developed countries other than those in the Middle East.



Figure 8. Examples of scenario terrains.

Human behavior cues were presented using 12 distinct virtual character models of identical height, similar size, and representing four skin tones: fair, light, medium, and dark. Since the theoretical foundation of the behavior cue analysis task domain is applicable in any cultural setting, the selected skin tones are representative of the diversity in human skin types as categorized by the Fitzpatrick Scale for skin pigmentation (Sachdeva, 2009). Virtual character skin tone was treated as a feature of the terrain. The fair, light, and dark models, including one male and one female per skin tone for a total of six different model types, appeared in training and practice scenarios portraying the Culturally Agnostic terrain. Middle Eastern terrain training and practice scenarios included only male models with medium skin tone. The female medium skin tone models provided by the VBS2 2.0 3D virtual model library all included garments that partially or completely veiled the faces and bodies of the models. The texture maps used to model the garments inhibited the proper depiction of animated behavior cues, therefore, the female medium skin tone models were unusable for this experiment. In order to provide the same number and variability of models as the Culturally Agnostic terrain, there were six different male medium skin tone model types included in the training and practice scenarios portraying the Middle Eastern terrain. All 12 models appeared in the pre-test and post-test scenarios as a means to control for potential effects of model type.

During scenarios, the virtual character models exhibited selected target or non-target behavior cue animations (Table 1). The Autodesk® MotionBuilder® 2011 3D animation software was used to create custom behavior cue animations of four target and two non-target behavior cues which were then imported into VBS2 2.0. The remaining two non-target behavior cues were selected from the existing VBS2 2.0 animations library.

Table 1.
Target and non-target behavior cues for experimental stimuli.

Behavior Cue		Description	Source	Classification
Target	Slap Hands	Back of one hand strikes the palm of the other hand repeatedly	Custom	Aggressiveness
	Clench Fists	Fingers are curled and squeezed into the palms	Custom	Aggressiveness
	Wring Hands	Alternate clasping, squeezing, and rubbing fingers and palm of one hand with the opposite hand repeatedly	Custom	Nervousness
	Check Six	Head turns to look over the shoulder followed by the body turning nearly 180°	Custom	Nervousness
Non-Target	Check “Watch”	Head angles down and one arm is raised slightly as if checking the time on a watch	Custom	N/A
	Rub Neck	Palm and fingers of one hand rubs the nape or side of the neck	Custom	N/A
	Idle Talking	Conversational behavior indicated by subtle hand and arm gestures	VBS2 2.0	N/A
	Cross Arms	Arms are bent at the elbows and overlap each other across the front of the body	VBS2 2.0	N/A

The virtual character models were arranged in tetrads with each position placed one meter from its adjacent position(s) and one meter from a central focal point (Figure 9). In each position, the virtual character models were angled to face the focal point of the tetrad. Distance from the route was set at two meters between the focal point of each tetrad and the center point of the UGS route. Consecutive tetrads alternated between left and right sides of the route and were placed at three meters between the focal points.

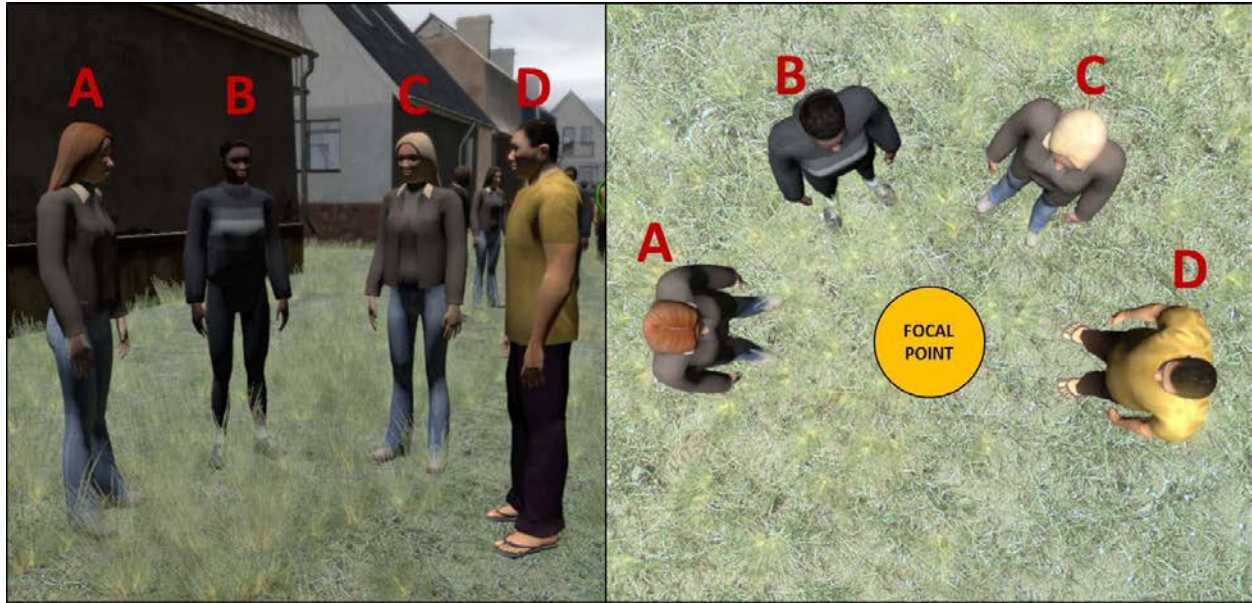


Figure 9. Example of virtual model positioning within a tetrad.

Virtual characters depicted only one behavior cue type at a time per position. The distribution of virtual character models and behavior cues was randomized and counterbalanced so that every model and cue combination was equally represented in each of the four tetrad positions. Tetrads contained either four non-target cues or one target cue with three non-target cues. During each scenario, behavior cue animations were triggered when the virtual UGS was 12 meters from the tetrad focal point, which is an acceptable distance for ISR ground operations (U.S. Army, 2007). Once triggered, animations looped repeatedly until all models in the tetrad were no longer in the UGS field-of-view. During pilot testing, synchronous animations within the same tetrad were determined to be distracting and less realistic. Therefore, upon triggering, the animation start times were randomly offset between a range of 68 to 840 milliseconds to limit the synchronicity.

Each instance that the behavior cue animations of a tetrad were triggered represented a single scenario event. With the virtual UGS speed of one and a half meters per second, the three

meter distance between tetrads, and the 12 meter animation triggering distance, the scenario event rate was approximately 30 events per minute. Theoretically, an event rate of 30 events per minute provides a moderate level of workload (Abich, Taylor, & Reinerman-Jones, 2013).

The experimental task required participants to monitor the simulated UGS surveillance footage, detect target behavior cues within each scenario event, and classify each detected target behavior cue as portraying aggressiveness or nervousness. The interface included two classification buttons labeled “Aggressiveness” and “Nervousness” (Figure 10). The experimental task procedure entailed using a computer mouse to click either classification button followed by clicking on the virtual character exhibiting the detected target behavior cue.

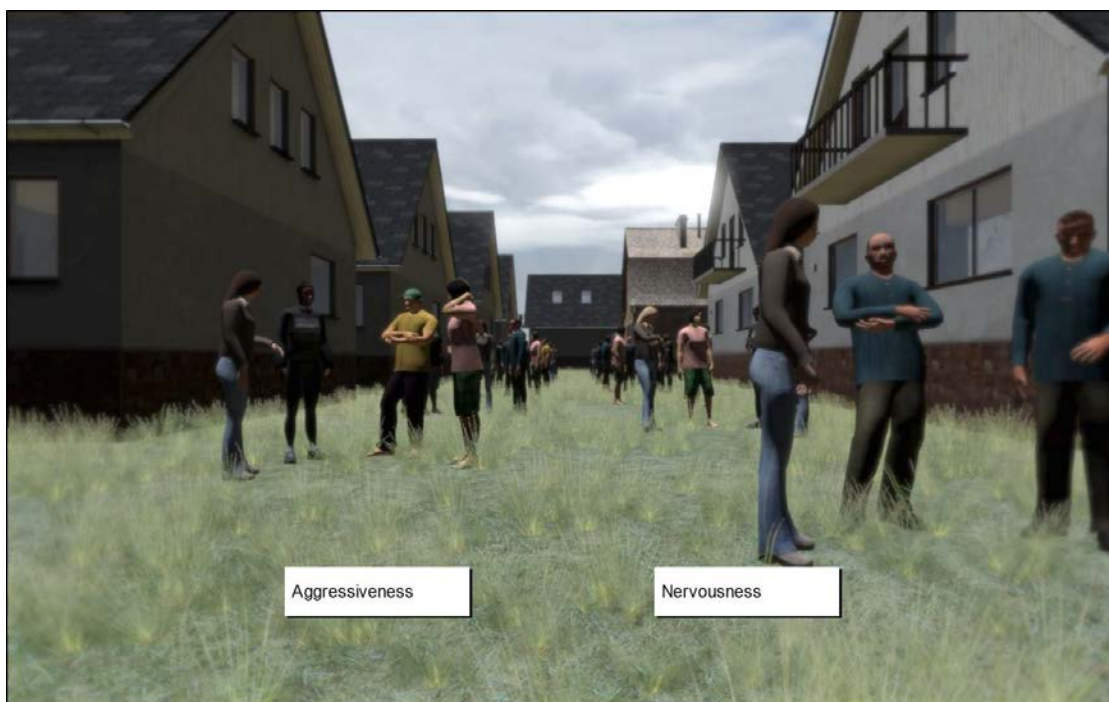


Figure 10. View of the VE display for experimental scenarios.

Experimental Design

This experiment followed a between groups design with one independent variable and three conditions. This experiment compared the effectiveness of two instructional strategies,

Massed Exposure and Highlighting, versus a Control condition in SBT for behavior cue analysis. Differences between the Massed Exposure and Highlighting conditions compared to the Control include variations in the number of target versus non-target events (Table 2) or the addition of an instructional support feature during training scenarios. Participants completed a total of two training scenarios, one each of the two scenario terrains (i.e., Middle Eastern and Culturally Agnostic) with corresponding virtual character models.

Across conditions, participants completed identical practice, pre-test, and post-test scenarios (Table 2). Participants completed a total of two practice scenarios, one of each scenario terrain with corresponding virtual character models. The pre-test and post-test scenarios depicted the Culturally Agnostic terrain and included all 12 virtual character models appearing in the training and practice scenarios.

Table 2.
Number of target versus non-target events per scenario type per condition.

	Control	Massed Exposure	Highlighting
Pre-Test Scenario (576 total events)	192 target events 384 non-target events	192 target events 384 non-target events	192 target events 384 non-target events
2 Training Scenarios (288 total events each)	96 target events 192 non-target events	192 target events 96 non-target events	96 highlighted target events 192 non-target events
2 Practice Scenarios (288 total events each)	96 target events 192 non-target events	96 target events 192 non-target events	96 target events 192 non-target events
Post-Test Scenario (576 total events)	192 target events 384 non-target events	192 target events 384 non-target events	192 target events 384 non-target events

Independent Variable

Control Condition

The Control condition represented a traditional SBT method. The training scenarios of the Control condition employed the baseline target event probability of one target event out of every three scenario events (Mogg & Bradley, 2002; Mogg & Bradley, 1999). Additionally, there were not any specific instructional supports included during training scenarios.

Massed Exposure Condition

The Massed Exposure condition followed the traditional SBT method of the Control, but presented a greater number of target events compared to the Control. During the Massed Exposure training scenarios, the target event probability was two target events out of every three events (Table 2).

Highlighting Condition

The Highlighting condition also followed the traditional SBT method of the Control with the addition of a non-content feature in the form of a translucent blue box overlaid on each virtual character model exhibiting target behavior cues (Figure 11). The target event probability was the same as the Control condition.



Figure 11. Example of a highlighted target.

Dependent Variables

Performance Metrics

Performance variables were collected during each scenario via a custom data logging program. Performance variables were calculated for overall performance and performance per target behavior cue type in the post-test scenario.

Detection Accuracy

Detection accuracy was calculated as the percentage of correctly detected targets out of the total number of targets presented during the scenario. Each model exhibiting a target behavior cue that the participant clicked was logged as a correct detection.

Classification Accuracy

Classification accuracy was calculated as the percentage of correctly classified targets out of the total number of targets presented during the scenario. If the selected classification button corresponded with the target behavior cue of the model clicked immediately following the button, then the response was logged as a correct classification.

Median Response Time

Response time was logged as the time, in milliseconds, between the moment a behavior cue animation was triggered and the time the participant clicked on the model exhibiting that cue. The value reported for data analysis was the median response time, which was calculated as the median response time value out of all the response times for only detected target behavior cues.

Percent Change

To assess the relative increase or decrease in performance scores from the pre-test to post-test scenarios, the percent change in detection accuracy, classification accuracy, and median response time was calculated for both overall performance and performance per target behavior cue type (Equation 1).

$$\text{Percent change} = \frac{X_{i\text{Post}} - X_{i\text{Pre}}}{X_{i\text{Pre}}} * 100 \quad (1)$$

Questionnaires

The following questionnaires were administered and logged using a customized software program. Participants responded to questionnaire items by selecting the numerical value for the desired response from a rating scale slider.

Engagement Measure

The Engagement Measure (APPENDIX A) assessed the participants' perceived level of engagement and involvement during the training scenarios. The Engagement Measure consists of seven items ($\alpha = .822$) selected from the engagement scale originally used by Charlton and Danforth (2005). Vocabulary regarding the specific experimental task was modified from the original scale to make the item verbiage relevant to the present experiment. Participants rated each item on a one to five scale, where one indicated strong disagreement and five indicated strong agreement with the statement. The engagement score was computed as the sum of the responses to each item, which included two reverse scored items.

Immersion Measure

The Immersion Measure (APPENDIX B) evaluated participants' sense of immersion during the training scenarios. The Immersion Measure includes eight items ($\alpha = .698$) selected from the immersion questionnaires created by Jennett, et al. (2008) for game-based experiences. The questionnaire items were tailored to the content of the present experiment by changing terms such *game* or *game events* to *scenario* or *scenario events*. Participants rated items along a one to five scale, where one indicated strong disagreement and five indicated strong agreement with the statement. The immersion score was computed as the sum of the responses to each item.

Presence Measure

The Presence Measure (APPENDIX C) measured participants' perceived level of presence experienced during interactions within the VE. The Presence Measure consists of 20 items (Total, $\alpha = .848$) selected from the original questionnaire developed by Witmer and Singer

(1998). Each item corresponds to one of four subscales including: Involvement and Control ($\alpha = .758$), Natural Interaction ($\alpha = .779$), Resolution ($\alpha = .684$), and Interface Quality ($\alpha = .677$).

Participants responded to items with a one to seven point scale to indicate low to high presence respectively. Scores are computed as the sum of items per subscale.

Procedure

Upon arrival to the laboratory space, the participant was instructed to turn off all cell phones and portable devices, remove watches or other time-pieces, and place them along with any other personal belongings in a secure area in the room away from the participant station. Then, the participant read and signed the Informed Consent form (APPENDIX D) which described the participant's rights as a study volunteer and the purpose, tasks, risks, and benefits of the experiment. Next, the participant verified that he or she fulfilled the study restrictions including a screening for color blindness using the Ishihara Test for Colour Blindness (Ishihara, 2013). All participants passed the color vision requirement. Next, the participant completed a paper-based demographics questionnaire (APPENDIX E) where he or she recorded information age, sex, highest level of education, military experience, current health state, and computer and video game proficiency. Then, the participant was randomly assigned to one of the three group conditions.

Next, the participant viewed a narrated slide presentation which familiarized him or her with the VE interface and the process to conduct the experimental task including monitoring the virtual UGS, selecting classification buttons, and clicking on detected targets. Following the slide presentation, the participant completed a familiarization scenario which allowed the participant to practice the target detection and classification procedure. To avoid priming effects,

the stimuli and classification categories during the familiarization scenario were unrelated to the stimuli appearing in the pre-test, post-test, training, and practice scenarios. The familiarization scenario stimuli consisted of colored barrels randomly placed along the route (Figure 12). Target barrels included six red and six yellow barrels; non-target barrels included eight brown, eight green, and eight white barrels. The participant detected as many target barrels as possible and classified each by color using the classification buttons labeled “Red” and “Yellow.” If the participant received a proficiency score of at least 75% detection accuracy, then he or she proceeded to the next phase of the experiment. Participants were provided up to two opportunities to achieve the minimum proficiency score. Four participants had to complete the familiarization scenario a second time. Only one participant did not fulfill proficiency requirements the second time and was dismissed.



Figure 12. Example view of the interface training scenario.

After becoming familiar with the detection and classification procedure, the participant viewed another narrated slide presentation describing the task for the following pre-test scenario. The pre-test scenario task required the participant to monitor the virtual UGS display and detect and classify targets that appeared to be exhibiting aggressiveness or nervousness based on his or her personal experience. The participant was permitted to take a five minute break to use the restroom and drink water after completing the pre-test scenario.

Next, the participant viewed a narrated, computer-based content training slide presentation that aligned with current Combat Profiling curriculum and ISR training doctrine (Colombo, Dolletski-Lazar, Coxe, & Tarr, 2012; U.S. Army, 2012a). The content training presented the purpose of behavior cue analysis in an ISR context, briefly explained the Kinesics domain, described the target behavior cues and the associated classifications, and provided example photographs of each target behavior cue. The photographs (APPENDIX F) included male and female models representing each of the four skin tones depicted within the VE. Models stood in front of a white background and their attire was standardized in order to minimize the risk of visual bias that may result from variations in color or style of clothing. The photographs were taken with a Canon EOS Rebel XTi 10.1 megapixel digital camera positioned 3.5 meters away from each model. Each target behavior cue was represented by a pair of poses to demonstrate the gesture movements. All poses were photographed from two angles, one with the model facing the camera and one with the model angled 45° to the model's right. These angles were similar to those of the virtual character models appearing within the scenarios. The photographs of each model, behavior cue, and pose angle was balanced throughout the presentation.

The content training slide presentation included a brief review where the participant viewed additional photographs depicting the target behavior cues and verbally classified each cue as exhibiting aggressiveness or nervousness. The participant's responses were recorded on a paper-based score sheet (APPENDIX G). After the review, the slide presentation provided a brief reminder of how to conduct the target detection and classification procedure within the VE. If the participant was assigned to either the Massed Exposure or Highlighting group conditions, a brief explanation of the instructional strategy applied to the following training scenarios was also included at the end of the slide presentation. The Massed Exposure condition slides stated that the training scenarios would present twice as many target events as opposed to non-target events. The Highlighting condition slides stated that during the training scenarios the targets would be highlighted by a translucent blue box, but the participant must still complete the detection and classification procedures. The Control condition did not include an instructional strategy explanation.

Following the content training slide presentation, the participant completed the two training scenarios, one depicting the Culturally Agnostic terrain and one depicting the Middle Eastern terrain. The order of the scenarios were randomized and counterbalanced. Upon completion of both training scenarios, the participant completed the Engagement Measure and the Immersion Measure. After completing the questionnaires, the participant was permitted another five minute break if needed to use the restroom and drink water.

The experiment resumed with the participant viewing a short, narrated slide presentation to introduce the practice scenarios to follow. The participant completed the two practice scenarios, one for each terrain. The order of the terrain types for the practice scenarios were the

same as the order of terrain types for the training scenarios. After the practice scenarios, the participant was offered one final five minute break to use the restroom and drink water if needed.

After the final break, the participant viewed a narrated slide presentation introducing the post-test scenario. The participant completed the post-test scenario followed by the Presence Measure. Finally, the participant was debriefed and dismissed. If the participant elected to receive monetary compensation, then he or she was provided a receipt signed by the experimenter. If the participant opted for course credit, then the experimenter notified the course professor after the participant was dismissed.

CHAPTER FOUR: RESULTS

Outliers

Upon review of frequency distributions, boxplots, and experimental log notes, two outliers were removed due to the participants' apparent lack of attentiveness during the task. This reduced both the Massed Exposure and Highlighting sample sizes by one. The revised sample sizes included: Control $n = 39$, Massed Exposure $n = 38$, and Highlighting $n = 39$.

Data Transformations

A constant of one was added to the detection accuracy, classification accuracy, and median response time scores for the pre-test and post-test. The purpose of this transformation was to avoid undefined values for the percent change variables due to a division by zero error.

Review of the frequency distributions for detection and classification accuracy scores revealed a substantial negative skew in each group condition. As a result, these distributions violated the assumption of normality. Therefore, to obtain a more normal distribution, scores were reflected to obtain a positive skew and a logarithmic transformation was applied. Equation 2 provides the transformation formula where k equals the largest value plus one.

$$\log_{10}(k - X_i) \quad (2)$$

Percent change variables also violated normality assumptions. Therefore, a logarithmic transformation was also applied to normalize the distributions. Due to several negative values, a large constant of 100 was added before logarithmic transformation (Equation 3).

$$\log_{10}(X_i + 100) \quad (3)$$

Analyses of Performance Metrics

Detection Accuracy

A one-way between-groups ANOVA revealed no significant difference in the effect of instructional strategy on overall detection accuracy, $F(2, 113) = .802, p = .451, \eta^2 = .014$.

Further, one-way between-groups ANOVAs revealed no significant differences in detection accuracy per target behavior cue type including: clenched fists, $F(2, 113) = .880, p = .417, \eta^2 = .015$, slapping hands, $F(2, 113) = .992, p = .374, \eta^2 = .017$, check six, $F(2, 113) = 2.15, p = .122, \eta^2 = .037$, and wringing hands, $F(2, 113) = .371, p = .691, \eta^2 = .007$.

A repeated measures ANOVA with Greenhouse-Geisser correction for sphericity in SPSS ($\epsilon = .935$) was conducted to assess the effect of target behavior cue types on detection accuracy. There was a significant effect of cue type on detection accuracy across conditions, $F(3, 113) = 463.93, p < .001, \eta^2 = 0.79$. Bonferroni pairwise comparisons revealed that the clench fists ($M = 56.77, SD = 23.06$), slapping hands ($M = 97.54, SD = 5.13$), check six ($M = 93.95, SD = 11.09$), and wringing hands ($M = 85.92, SD = 14.49$) cues differed significantly from each other at a significance level of $p < .001$ with clenched fists having the lowest detection accuracy scores (Figure 13).

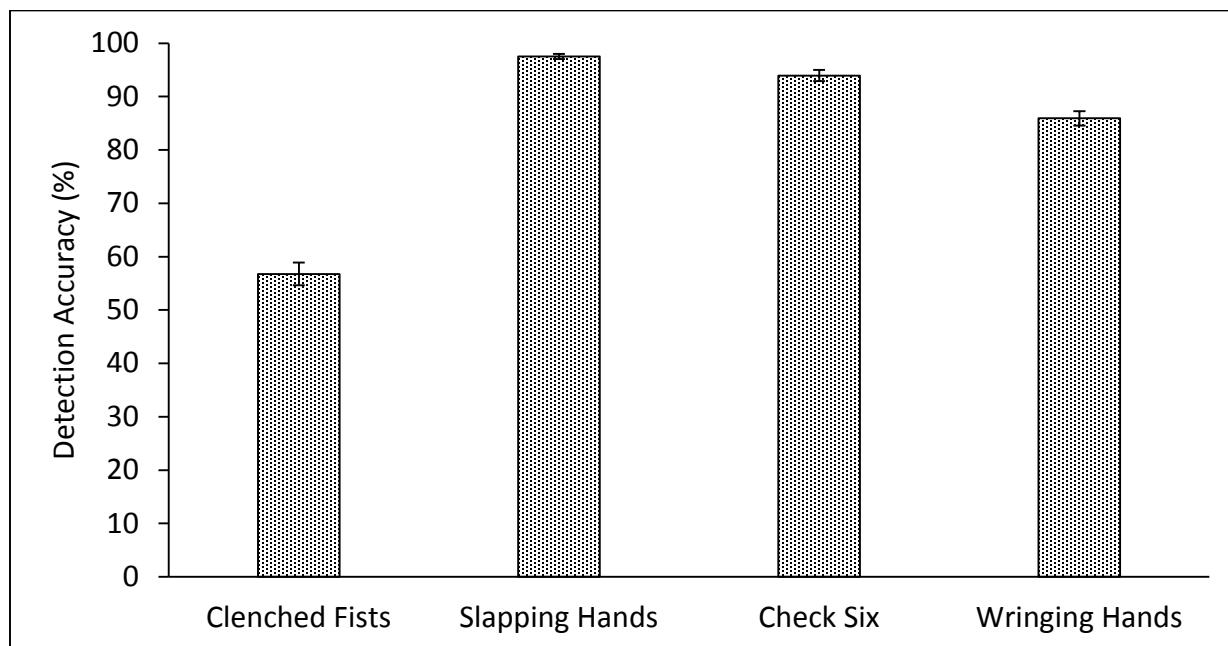


Figure 13. Mean detection accuracy across conditions per target behavior cue type. Error bars indicate standard error.

Classification Accuracy

A one-way between-groups ANOVA revealed no significant difference in the effect of instructional strategy on overall classification accuracy scores, $F(2, 113) = 1.09, p = .338, \eta^2 = .019$. Additionally, one-way between-groups ANOVAs revealed no significant effect of instructional strategy on classification accuracy per target behavior cue type including: clenched fists, $F(2, 113) = 1.04, p = .356, \eta^2 = .018$, slapping hands, $F(2, 113) = 1.69, p = .189, \eta^2 = .02$, check six, $F(2, 113) = 2.21, p = .115, \eta^2 = .038$, and wringing hands, $F(2, 113) = .391, p = .677, \eta^2 = .007$.

A repeated measures ANOVA with Greenhouse-Geisser correction for sphericity in SPSS ($\epsilon = .866$) was conducted to assess the effect of target behavior cue type on classification accuracy across conditions. There was a significant effect of cue type on classification accuracy across conditions, $F(3, 113) = 220.19, p < .001, \eta^2 = .72$. At a significance level of $p < .001$,

Bonferroni pairwise comparisons indicated that classification accuracy scores for the clenched fists ($M = 52.35$, $SD = 23.22$), slapping hands ($M = 92.19$, $SD = 11.45$), check six ($M = 89.78$, $SD = 15.19$), and wringing hands ($M = 81.16$, $SD = 16.49$) cues differed significantly from each (Figure 14).

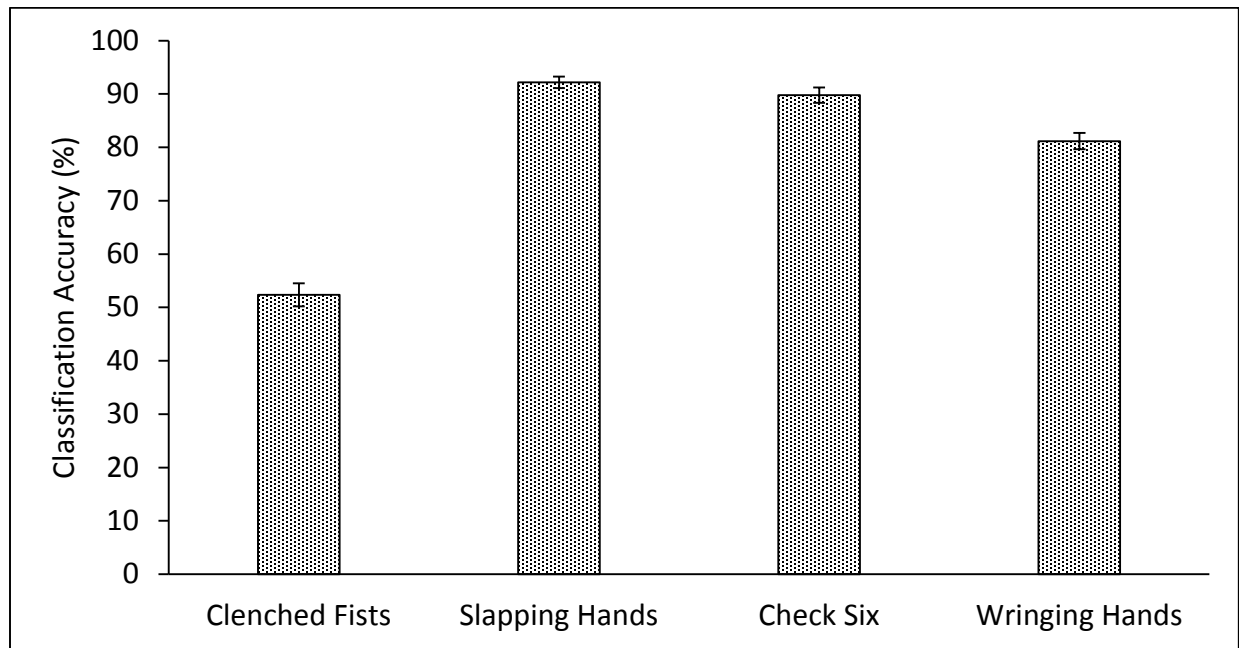


Figure 14. Mean classification accuracy across conditions per target behavior cue type. Error bars indicate standard error.

Median Response Time

One-way between-groups ANOVAs were conducted to assess the impact of instructional strategy on median response time. While there was no significant difference in overall median response time, $F(2, 113) = 2.93$, $p = .058$, $\eta^2 = .049$, between groups, there was a significant effect of instructional strategy on median response time for target behavior cues in the aggressiveness classification. There was a significant effect of instructional strategy on clenched fists median response time, $F(2, 113) = 4.36$, $p = .015$, $\eta^2 = .071$. Post-hoc tests using the Bonferroni correction revealed that participants in the Massed Exposure group ($M = 6.69$, $SD =$

.807) correctly responded to the clenched fists cue significantly faster than participants in the Control group ($M = 7.23$, $SD = .746$), $p = .013$ (Figure 15). However, median response time in the Highlighting group did not differ significantly for the clenched fists cue ($M = 6.89$, $SD = .847$).

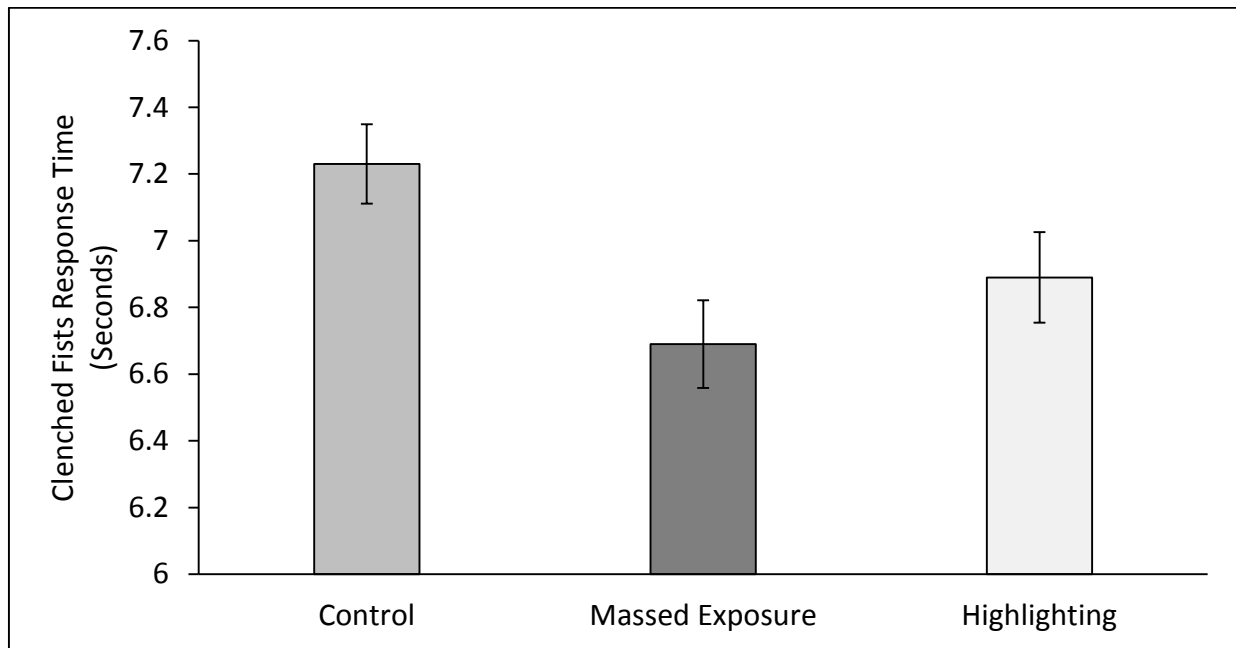


Figure 15. Median response time means for the clenched fists cue. Error bars indicate standard error.

Likewise, there was a significant effect of instructional strategy on slapping hands median response time, $F(2, 113) = 3.44$, $p = .035$, $\eta^2 = .057$. Post-hoc tests using Bonferroni corrections indicated that participants responded to slapping hands cues faster in the Massed Exposure group ($M = 5.36$, $SD = .547$) than those in the Control group ($M = 5.71$, $SD = .675$), $p = .043$ (Figure 16). Median response time in the Highlighting group did not differ significantly for the slapping hands cue ($M = 5.43$, $SD = .587$).

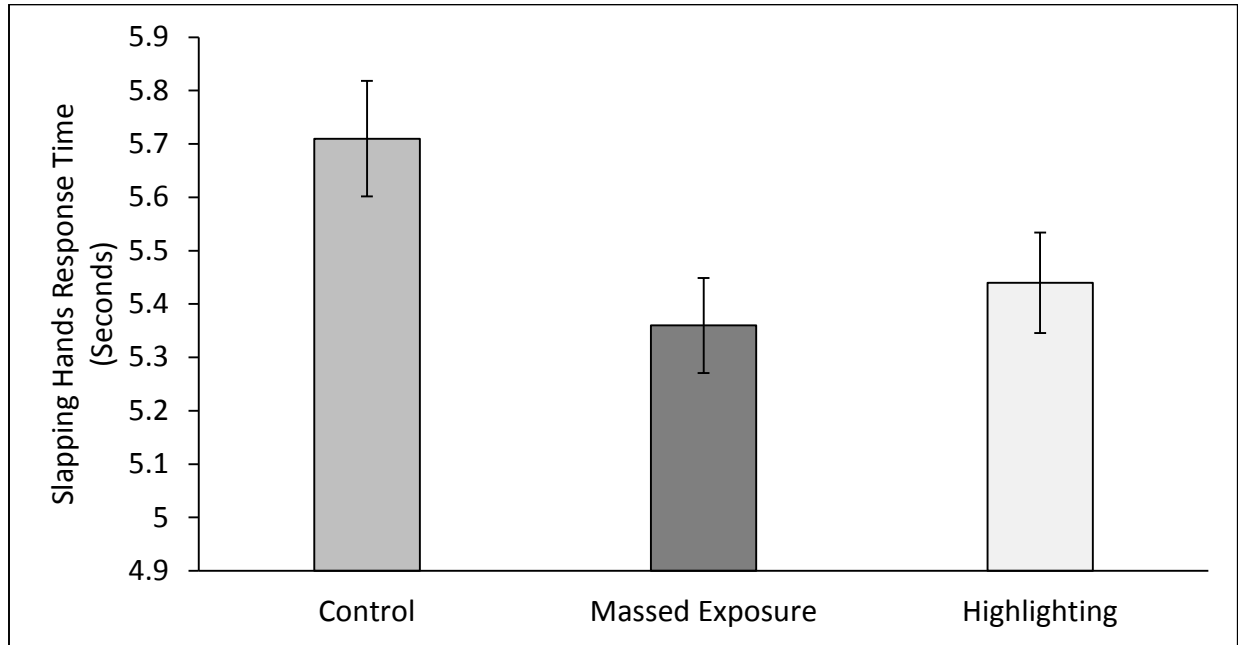


Figure 16. Median response time means for the slapping hands cue. Error bars indicate standard error.

Median response times of the target behavior cues in the nervousness classification did not differ significantly between groups for check six, $F(2, 113) = 2.91, p = .058, \eta^2 = .049$, and wringing hands, $F(2, 113) = 2.58, p = .080, \eta^2 = .044$.

A repeated measures ANOVA with Greenhouse-Geisser correction for sphericity in SPSS ($\epsilon = .866$) was conducted to assess the effect of the target behavior cue types on median response time. A significant effect of cue type on median response time was revealed across conditions, $F(3, 113) = 220.19, p < .001, \eta^2 = .83$. Bonferroni pairwise comparisons indicated a significant difference between cue types at the $p < .05$ level for median response time including: clench fists ($M = 6.94, SD = .824$), slapping hands ($M = 5.50, SD = .619$), check six ($M = 5.40, SD = .722$), and wringing hands ($M = 5.89, SD = .689$). Clenched fists had the highest median response time indicating that it took longer for participants to detect this cue compared to other cue types (Figure 17).

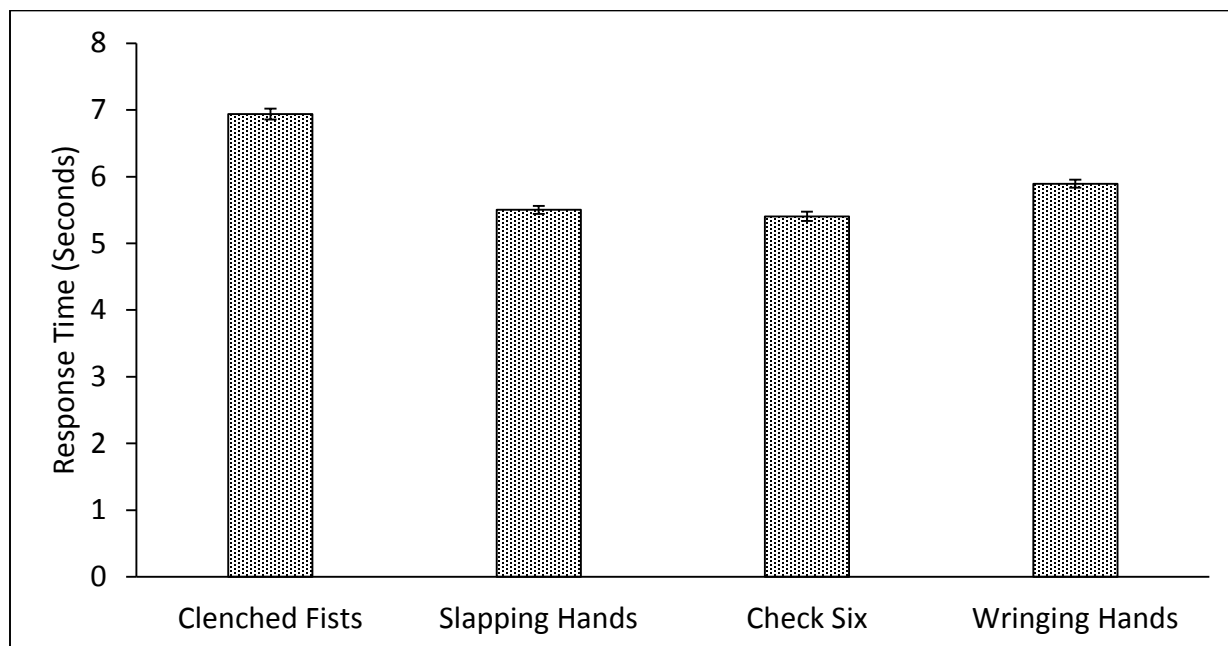


Figure 17. Median response time means across conditions per target behavior cue type. Error bars indicate standard error.

Percent Change

The impact of instructional strategy on the percent change in performance scores (i.e., detection accuracy, classification accuracy, and median response time) from the pre-test scenario to the post-test scenario was assessed with a series of one-way between-groups ANOVAs. There was not a significant effect of instructional strategy on the percent change for overall detection accuracy, $F(2, 113) = .018, p = .982, \eta^2 < .001$, overall classification accuracy, $F(2, 113) = .144, p = .866, \eta^2 = .003$, nor overall median response time, $F(2, 113) = 2.15, p = .122, \eta^2 = .039$. Percent change variables also did not differ significantly per cue type.

A series of repeated measures ANOVAs revealed a significant effect of cue type on the percent change of each performance metric. A significant effect of cue type on the percent change for detection accuracy, with Greenhouse-Geisser correction in SPSS ($\epsilon = .883$), was found across conditions, $F(2.68, 308.18) = 48.56, p < .001, \eta^2 = .30$. Bonferroni pairwise

comparisons indicated that clenched fists cue ($M = 1930.64$, $SD = 2746.67$) had a significantly greater increase in detection accuracy at the $p < .001$ level compared to slapping hands ($M = 178.93$, $SD = 1009.31$), check six ($M = 1003.44$, $SD = 2243.58$), and wringing hands ($M = 802.09$, $SD = 2018.77$). The increase in detection accuracy was significantly lower for the slapping hands cue compared to the check six and wringing hands cues, $p < .001$. The check six and wringing hands cues did not differ significantly for the percent change in detection accuracy, $p = .072$. Figure 18 illustrates the average detection accuracy percent change between target behavior cue types.

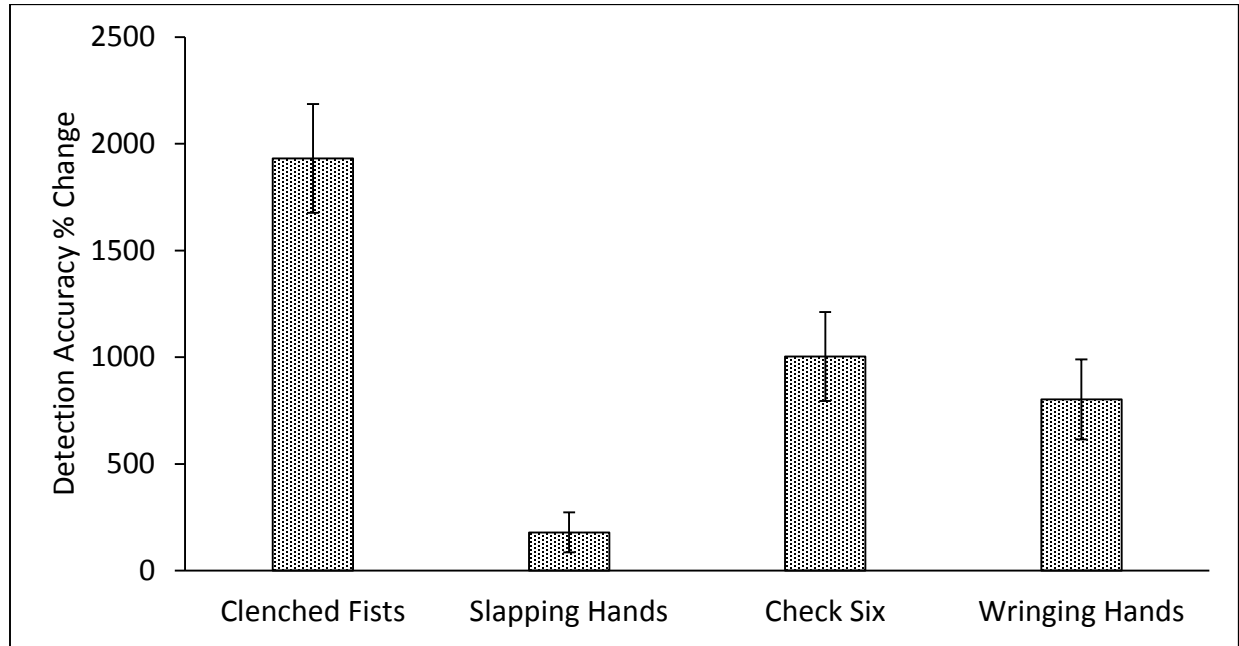


Figure 18. Percent change in detection accuracy per target behavior cue type across conditions. Error bars indicate standard error.

Cue type also had a significant effect on the percent change for classification accuracy with Greenhouse-Geisser correction in SPSS ($\epsilon = .893$), $F(2.73, 314.38) = 46.83$, $p < .001$, $\eta^2 = .48$. Bonferroni pairwise comparisons indicated that the increase in classification accuracy for the slapping hands cue ($M = 244.21$, $SD = 1310.42$) was significantly less at the $p < .001$ level than

the clenched fists ($M = 2316.42$, $SD = 327.52$), check six ($M = 5063.04$, $SD = 3996.48$), and wringing hands ($M = 2713.07$, $SD = 3499.19$) cues. The clenched fists and wringing hands cues did not differ significantly, $p = 1.00$. Figure 19 illustrates the differences in the average classification accuracy percent change between target behavior cue types.

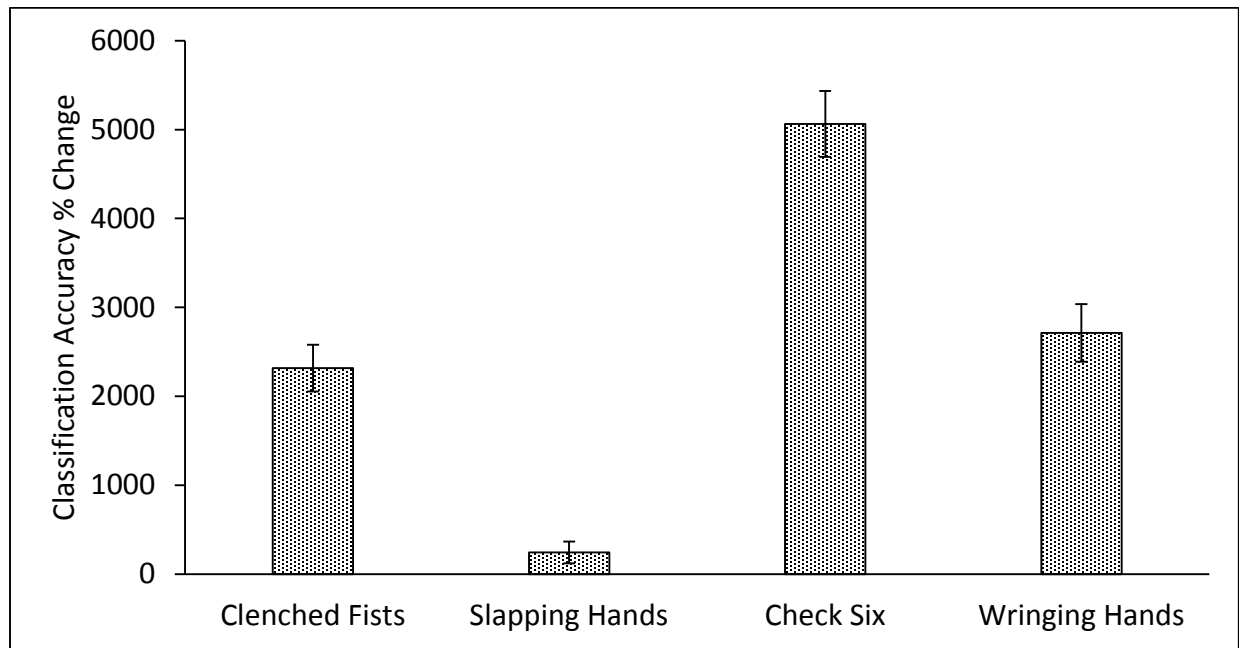


Figure 19. Percent change in classification accuracy per target behavior cue type across conditions. Error bars indicate standard error.

Finally, a significant effect of cue type also was also found for the percent change in median response time with Greenhouse-Geisser correction in SPSS ($\epsilon = .606$), $F(1.82, 209) = 36.56$, $p < .001$, $\eta^2 = .24$. Bonferroni pairwise comparisons revealed that the increase in median response time for the clenched fists cue ($M = 201.42$, $SD = 327.52$) was significantly different at the $p < .001$ level from the slapping hands ($M = -9.04$, $SD = 52.21$), check six ($M = 15.99$, $SD = 126.89$), and wringing hands ($M = 21.53$, $SD = 142.51$) cues. The slapping hands, check six, and wringing hands cues did not differ significantly from each other (Figure 20).

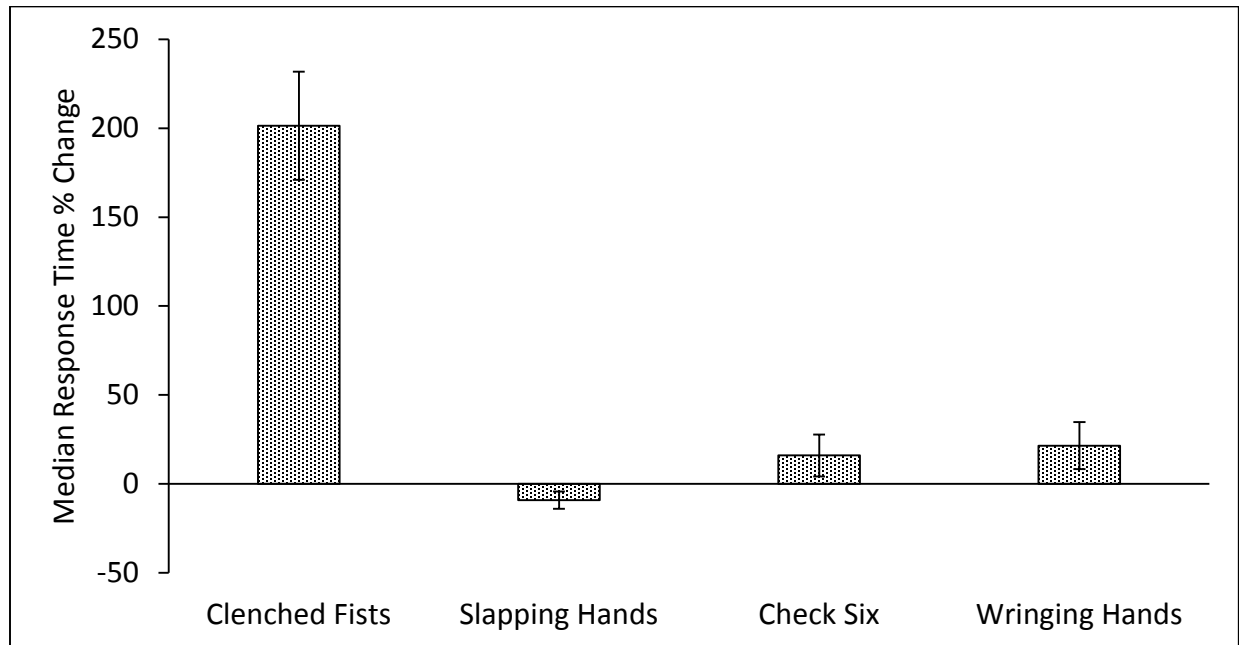


Figure 20. Percent change in median response time per target behavior cue type across conditions. Error bars indicate standard error.

Analyses of Questionnaires

Engagement Measure

A one-way between-groups ANOVA revealed no significant difference in the effect of instructional strategy on the level of engagement as assessed by the Engagement Measure, $F(2, 113) = 2.03, p = .136, \eta^2 = .035$. However, a series of one-way between-groups ANOVAs were conducted to compare the ratings per survey item. There was a significant difference in ratings for survey item seven, *I like the challenge that using a virtual environment for behavior cue detection training provided*, $F(2, 113) = 4.78, p = .010, \eta^2 = .008$. Post-hoc tests with Bonferroni correction revealed a significantly lower rating in the Highlighting group ($M = 2.95, SD = .887$) versus the Control group ($M = 3.59, SD = 1.04$), $p = .011$. This difference indicates that the preferred level of challenge during training scenarios occurred in the Control condition. The

rating on survey item seven in the Massed Exposure group ($M = 3.11$, $SD = .924$) did not differ significantly, but was still less than the Control group. Figure 21 illustrates the results for survey item seven of the Engagement Measure.

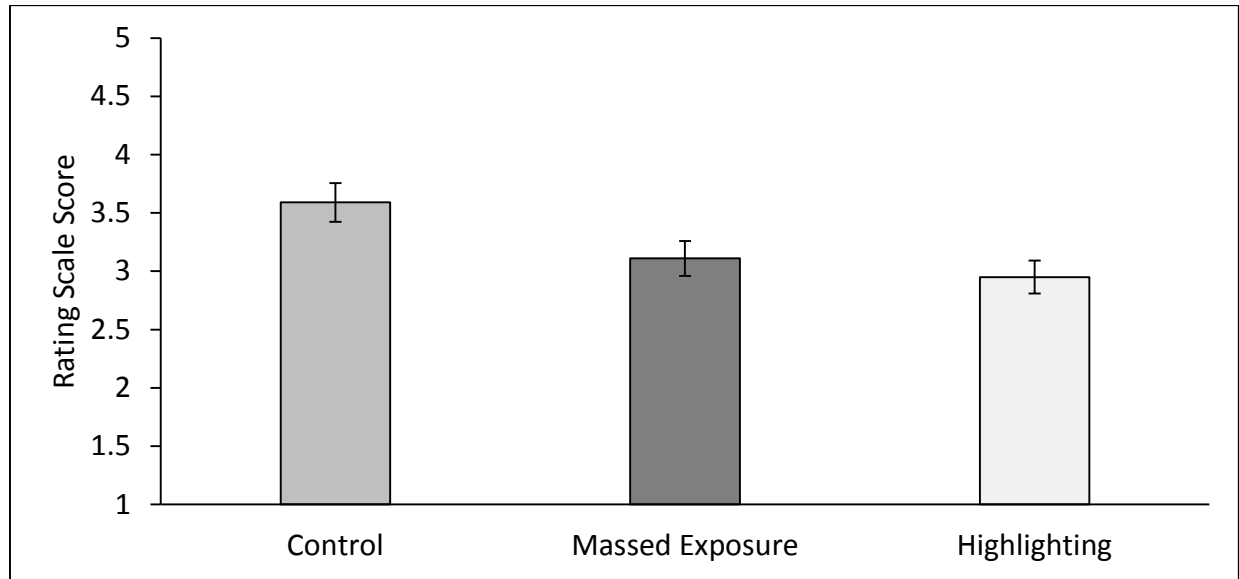


Figure 21. Engagement Measure means between conditions for item seven, *I like the challenge that using a virtual environment for behavior cue detection training provided*. Error bars indicate standard error.

Immersion Measure

A one-way between-groups ANOVA revealed no significant difference in the effect of instructional strategy on the level of immersion as assessed by the Immersion Measure, $F(2, 113) = 1.71$, $p = .185$, $\eta^2 = .029$. Another series of one-way between-groups ANOVAs were conducted to compare the ratings per survey item revealing a significant difference in ratings for survey item eight, *The scenarios were challenging*, $F(2, 113) = 8.13$, $p = .001$, $\eta^2 = .013$. Bonferroni correction post-hoc tests revealed a significantly lower rating in the Highlighting group ($M = 2.05$, $SD = .857$) versus the Control group ($M = 2.87$, $SD = 1.11$), $p = .002$, and the Massed Exposure ($M = 2.87$, $SD = 1.12$), $p = .002$. This indicates that the training scenarios in

the Highlighting condition were perceived as less challenging than training scenarios in the Control and Massed Exposure conditions as illustrated in Figure 22.

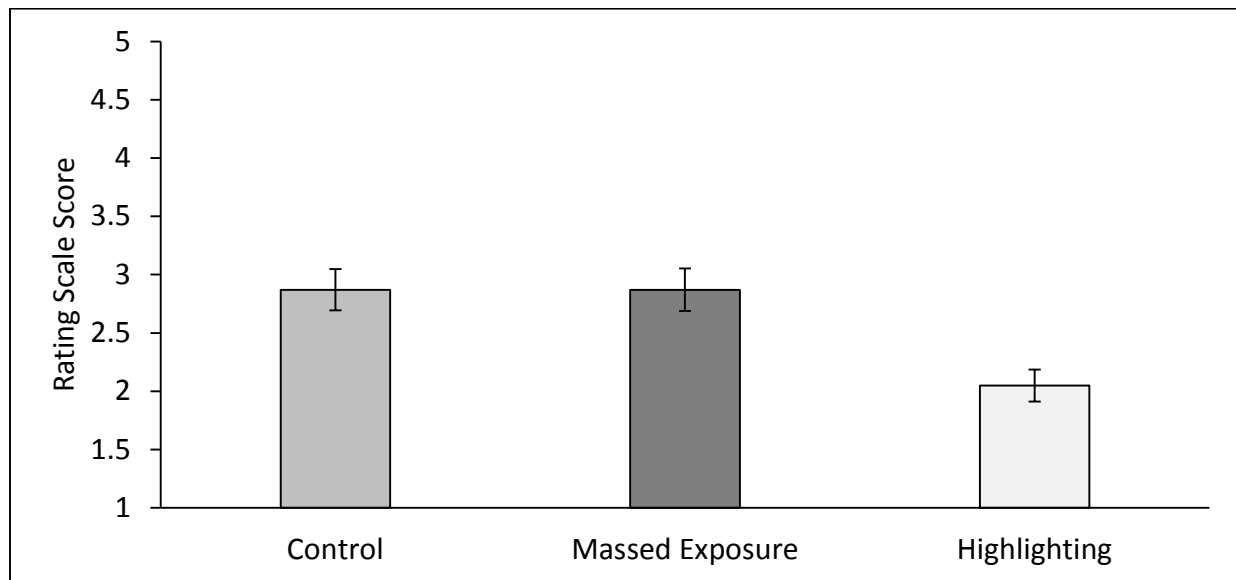


Figure 22. Immersion Measure means between conditions for item 8, *The scenarios were challenging*. Error bars indicate standard error.

Presence Measure

A one-way between-groups ANOVA revealed no significant difference in the effect of instructional strategy on the level of total presence as assessed by the sum of items on the Presence Measure, $F(2, 113) = .117, p = .890, \eta^2 = .002$. Further, there were no significant differences between instructional strategy groups for each Presence Measure subscales including: Involvement and Control, $F(2, 113) = .113, p = .893, \eta^2 = .002$, Natural Interaction, $F(2, 113) = .162, p = .850, \eta^2 = .003$, Resolution, $F(2, 113) = .647, p = .526, \eta^2 = .011$, and Interface Quality, $F(2, 113) = .852, p = .429, \eta^2 = .015$. There were also no significant differences between conditions in the ratings for individual survey items.

Correlates of Performance Metrics and Questionnaires

For each condition, a series of Pearson's r correlations were conducted to assess the relationships between performance and questionnaire scores. Due to the reverse scoring of the detection and classification accuracy variables prior to log transformation, the signs indicating the correlation direction for these variables have been reversed.

Control Group Correlates

There were significant, moderately positive correlations between detection accuracy and engagement, measured by the Engagement Measure, and between detection accuracy and immersion, measured by the Immersion Measure. Per target behavior cue, only the clenched fists cue had a significant, moderately positive correlation between detection accuracy and engagement and between detection accuracy and immersion. The significant relationships indicate that greater detection accuracy scores were correlated with higher engagement or immersion ratings both overall and for the clenched fists cue. Table 3 lists the correlation results for detection accuracy, engagement, and immersion for the Control group.

Table 3.
Control group correlates between detection accuracy, engagement, and immersion.

	Engagement Measure	Immersion Measure
Performance Metrics	r	r
Overall Detection Accuracy	.319*	.358*
Clenched Fist Detection Accuracy	.334*	.442**
Slapping Hands Detection Accuracy	.153	.215
Check Six Detection Accuracy	.286	.171
Wringing Hands Detection Accuracy	.146	.137

Note. Correlation directions have been reversed due to data transformation.

* $p \leq .05$; ** $p \leq .01$

Classification accuracy was positively correlated to both engagement and immersion. Significant, moderately positive correlations between classification accuracy and engagement were also revealed for the clenched fists, check six, and wringing hands cues. Between classification accuracy and immersion, only the clenched fists and wringing hands cues had significant, moderately positive correlations. Similar to the preceding detection accuracy correlations, these significant relationships indicate that greater classification accuracy scores were correlated with higher engagement and immersion ratings both overall and for select cues. Table 4 lists the correlation results for classification accuracy, engagement, and immersion for the Control group.

Table 4.
Control group correlates between classification accuracy, engagement, and immersion.

	Engagement Measure	Immersion Measure
Performance Metrics	<i>r</i>	<i>r</i>
Overall Classification Accuracy	.400*	.386*
Clenched Fist Classification Accuracy	.377*	.447**
Slapping Hands Classification Accuracy	.310	.271
Check Six Classification Accuracy	.325*	.245
Wringing Hands Classification Accuracy	.345*	.326*

Note. Correlation directions have been reversed due to data transformation.

* $p \leq .05$; ** $p \leq .01$

Neither engagement nor immersion were significantly correlated to overall and per cue median response time. Additionally, engagement was not significantly correlated to any percent change performance variables. However, the percent change in detection accuracy and percent change in classification accuracy were both significantly correlated to immersion. Per cue, percent change in detection accuracy and percent change in median response time were

significantly correlated to immersion for the wringing hands cue only. These positive relationships indicate that higher immersion ratings were correlated to increases in overall detection and classification accuracy, as well as, increased wringing hands response time. Table 5 lists the significant overall and wringing hands correlations between the percent change performance scores and immersion for the Control group.

Table 5.
Control group correlates between percent change performance metrics and immersion.

Performance Metrics	Immersion Measure
	<i>r</i>
Overall % Change Detection Accuracy	.353*
Overall % Change Classification Accuracy	.329*
Wringing Hands % Change Detection Accuracy	.319*
Wringing Hands % Change Median Response Time	.381*

* $p \leq .05$; ** $p \leq .01$

There were several significant correlations between presence, assessed by the Presence Measure, and performance. Overall detection accuracy was positively correlated to total presence in addition to the Involvement and Control, Natural Interaction, and Interface Quality subscales. The Resolution subscale did not correlate significantly to detection accuracy. These relationships were also consistent for several of the target behavior cue types where higher presence or subscales of presence were correlated to greater detection accuracy performance. Notably, the correlations between clenched fists detection accuracy and presence were similar in strength and direction to the correlations between overall detection accuracy and presence. Table 6 provides a list of the presence and detection accuracy correlations for the Control group.

Table 6.

Control group correlates between detection accuracy and presence.

	Total Presence	Involvement and Control	Natural Interaction	Resolution	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Overall Detection Accuracy	.505**	.448**	.338*	.189	.391*
Clenched Fists Detection Accuracy	.485**	.410**	.397*	.186	.348*
Slapping Hands Detection Accuracy	.363*	.265	.225	.259	.381*
Check Six Detection Accuracy	.369*	.390*	.131	-.045	.377*
Wringing Hands Detection Accuracy	.394*	.377*	.207	.173	.270

Note. Correlation directions have been reversed due to data transformation.

* $p \leq .05$; ** $p \leq .01$

Overall classification accuracy also positively correlated to total presence and the Involvement and Control, Natural Interaction, and Interface Quality subscales. There was not a significant correlation between the Resolution subscale and classification accuracy. Correlations between classification accuracy and presence per target behavior cue type also revealed several significant relationships where higher presence or subscales of presence were correlated to greater classification accuracy performance. The strength and direction of the correlations between classification accuracy and presence for the clenched fists and wringing hands cues were similar to those for overall classification accuracy. The correlations between classification accuracy and presence are listed in Table 7.

Table 7.

Control group correlates between classification accuracy and presence.

	Total Presence	Involvement and Control	Natural Interaction	Resolution	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Overall Classification Accuracy	.505**	.424**	.369*	.160	.456**
Clenched Fists Classification Accuracy	.478**	.400*	.389*	.165	.374*
Slapping Hands Classification Accuracy	.342*	.229	.286	.118	.432**
Check Six Classification Accuracy	.429**	.403*	.280	.120	.305
Wringing Hands Classification Accuracy	.495**	.426**	.326*	.167	.449**

Note. Correlation directions have been reversed due to data transformation.

* $p \leq .05$; ** $p \leq .01$

Total presence, Resolution, and Interface Quality scores on the Presence Measure were each positively correlated to the pre-test to post-test percent change for both detection and classification accuracy (Table 8). Therefore, increases in overall accuracy performance were moderately correlated with higher presence ratings.

Table 8.

Control group correlates between overall percent change in detection and classification accuracy and presence.

	Total Presence	Resolution	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>	<i>r</i>
Overall % Change Detection Accuracy	.292	.356*	.385*
Overall % Change Classification Accuracy	.354*	.352*	.356*

* $p \leq .05$; ** $p \leq .01$

Per target behavior cue, only the check six and wringing hands cues revealed significant correlations between presence and percent change metrics. For the check six cue, increases in detection accuracy performance correlated with higher ratings on the Interface Quality subscale. Increases in check six median response time also correlated to higher ratings for total presence and Interface Quality. The percent change in check six classification accuracy did not significantly correlate to presence. Table 9 lists select correlation results between the check six percent change performance metrics and presence scores.

Table 9.
Control group correlates between check six percent change metrics and presence.

	Total Presence	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>
Check Six % Change Detection Accuracy	.292	.371*
Check Six % Change Classification Accuracy	.315	.208
Check Six % Change Median Response Time	.318*	.397*

* $p \leq .05$; ** $p \leq .01$

For the wringing hands cue, increases in both detection accuracy and median response time were moderately correlated to higher Interface Quality scores. An increase in wringing hands classification accuracy was also positively correlated to higher total presence and Involvement and Control scores. Table 10 lists selected correlations between the percent change performance metrics and presence for the wringing hands cue.

Table 10.

Control group correlates between wringing hands percent change metrics and presence.

	Total Presence	Involvement and Control	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>	<i>r</i>
Wringing Hands % Change Detection Accuracy	.193	.123	.355*
Wringing Hands % Change Classification Accuracy	.339*	.325*	.285
Wringing Hands % Change Median Response Time	.283	.235	.410**

* $p \leq .05$; ** $p \leq .01$

Collectively, the Control group correlations described above indicate that higher engagement, immersion, or presence ratings moderately correlated to better detection and classification accuracy. Furthermore, these relationships reveal that higher engagement, immersion, and presence scores were correlated to percent change increases in performance. However, the results were not universally consistent across target behavior cue types.

Massed Exposure Group Correlates

The Massed Exposure condition had several significant correlations between performance metrics and questionnaire variables that were uniquely distinct compared to the Control and Highlighting groups. There was only one significant correlation between performance metrics and engagement; greater check six detection accuracy was positively correlated with higher engagement, $r = .357$, $n = 38$, $p = .028$. Note that the reported correlation direction has been reversed to account for data transformation. Immersion did not significantly correlate with any performance metrics for the Massed Exposure condition.

The remaining significant correlations for Massed Exposure were between performance and presence. There was a positive correlation between greater overall detection accuracy scores

and higher ratings on the Resolution subscale of the Presence Measure, $r = .322$, $n = 38$, $p = .049$. Also, greater wringing hands detection accuracy positively correlated with higher ratings on the Interface Quality subscale of the Presence Measure, $r = .364$, $n = 38$, $p = .025$. Once again, the direction of these correlations have been reversed to account for data transformation.

Correlations between Massed Exposure percent change performance variables and presence revealed moderately negative relationships only in select cases. Increased percent change in detection accuracy correlated to lower total presence ratings. This relationship also held for the slapping hands and wringing hands cues only. Increased detection accuracy for the slapping hands cue was also correlated with lower scores on the Involvement and Control subscale, while increased wringing hands detection accuracy was correlated with lower Natural Interaction scores. Table 11 lists the correlations between percent change in detection accuracy and presence for the Massed Exposure condition.

Table 11.
Massed Exposure correlates between percent change in detection accuracy and presence.

	Total Presence	Involvement and Control	Natural Interaction	Resolution	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Overall % Change Detection Accuracy	-.320*	-.286	-.221	-.078	-.244
Clenched Fists % Change Detection Accuracy	.037	.065	-.015	.261	-.158
Slapping Hands % Change Detection Accuracy	-.333*	-.337*	-.099	-.173	-.235
Check Six % Change Detection Accuracy	.053	.056	.015	.077	.003
Wringing Hands % Change Detection Accuracy	-.338*	-.272	-.353*	-.015	-.251

* $p \leq .05$; ** $p \leq .01$

Percent change in classification accuracy and presence correlations revealed some consistency with the preceding percent change in detection accuracy results. Increased percent change in classification accuracy negatively correlated to lower total presence ratings, in addition to lower Involvement and Control subscale ratings. These significant relationships held for the slapping hands percent change in classification accuracy. Two significant positive correlations emerged between increased check six classification accuracy and higher Interface Quality and increased wringing hands classification accuracy and higher Resolution scores. Table 12 lists the correlations between percent change in classification accuracy metrics and presence for the Massed Exposure condition.

Table 12.
Massed Exposure correlates between percent change in classification accuracy and presence.

	Total Presence	Involvement and Control	Natural Interaction	Resolution	Interface Quality
Performance Metrics	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Overall % Change Classification Accuracy	-.385*	-.373*	-.284	-.101	-.197
Clenched Fists % Change Classification Accuracy	-.047	-.019	-.054	.128	-.158
Slapping Hands % Change Classification Accuracy	-.343*	-.345*	-.115	-.180	-.233
Check Six % Change Classification Accuracy	-.012	-.106	-.136	-.014	.380*
Wringing Hands % Change Classification Accuracy	-.088	-.061	-.255	.328*	-.118

* $p \leq .05$; ** $p \leq .01$

Finally, there were only two significant correlations between the percent change in median response time metrics and presence. Decreased clenched fists median response time was

moderately correlated with higher ratings on the Interface Quality subscale, $r = -.367$, $n = 38$, $p = .023$. Conversely, decreased slapping hands median response time was also moderately correlated with lower Involvement and Control ratings, $r = .322$, $n = 38$, $p = .048$.

Highlighting Group Correlates

Although the Highlighting group had the fewest significant correlations of all the conditions, the relationships that emerged were exclusive to the Highlighting condition only. Further, only performance variables for the clenched fists cue revealed significant relationships with engagement and immersion scores. Engagement was negatively correlated with the percent change in detection accuracy and the percent change in median response time for the clenched fists cue. Likewise, immersion was negatively correlated with the percent change in detection accuracy and the percent change in median response time for the clenched fists cue. Altogether, these relationships indicate that greater engagement and immersion scores were moderately correlated with decreases in detection accuracy and median response time of the clenched fists cue. Table 13 provides a list of the significant correlation results between performance and questionnaire variables in the Highlighting condition.

Table 13.

Highlighting group correlates between percent change performance metrics, engagement and immersion.

	Engagement Measure	Immersion Measure
Performance Metrics	<i>r</i>	<i>r</i>
Clenched Fists % Change Detection Accuracy	-.339*	-.318*
Clenched Fist % Change Median Response Time	-.479**	-.368*

* $p \leq .05$; ** $p \leq .01$

Regressions

A series of multiple regression analyses were conducted to assess if participants' ratings on the questionnaires significantly predicted performance. Two sets of predictor variables were tested for each instructional strategy condition. The first set of predictors included total scores on the measures of engagement, immersion, and presence. The second set of predictors consisted of the total engagement and immersion scores, and the individual subscale scores of the Presence Measure.

Control Group Regressions

The majority of all significant regression models emerged in the Control group. The first set of predictors, including engagement, immersion and total presence, explained approximately 26% of the variance in overall detection accuracy, $R^2 = .263$, $F(3, 35) = 4.16$, $p = .013$. It was found that the total presence score was the significant contributor to overall detection accuracy, $\beta = .442$, $p = .021$. Between target behavior cue types, the engagement, immersion, and total presence predictors explained approximately 27% of the variance in detection accuracy for the clenched fists cue only, $R^2 = .271$, $F(3, 35) = 4.33$, $p = .011$, however, none of the individual predictors were significant.

The engagement, immersion and total presence predictors explained nearly 29% of the variance in overall classification accuracy, $R^2 = .288$, $F(3, 35) = 4.71$, $p = .007$, and indicated total presence as the significant contributor, $\beta = .401$, $p = .033$. Between target behavior cue types, this relationship held for clenched fists, $R^2 = .274$, $F(3, 35) = 4.39$, $p = .010$, check six, $R^2 = .213$, $F(3, 35) = 3.15$, $p = .037$, and wringing hands, $R^2 = .261$, $F(3, 35) = 4.12$, $p = .013$.

However, total presence was the significant contributor to classification accuracy for only check six, $\beta = .410$, $p = .037$, and wringing hands, $\beta = .446$, $p = .020$.

The final significant regression model including the first set of predictor variables indicated that engagement, immersion, and total presence explained nearly 21% of the variance in the percent change in detection accuracy for the slapping hands cue only, $R^2 = .206$, $F(3, 35) = 3.02$, $p = .043$. The greatest significant contributor was engagement, $\beta = -.556$, $p = .009$, and the secondary significant contributor was immersion, $\beta = .534$, $p = .021$.

The second set of predictor variables, including engagement, immersion, and the presence subscales, explained approximately 31% of the variance in check six detection accuracy, $R^2 = .315$, $F(6, 32) = 2.45$, $p = .046$. The Involvement and Control subscale was the greatest significant contributor, $\beta = .440$, $p = .041$. The Interface Quality subscale was the secondary significant contributor, $\beta = .379$, $p = .026$.

The engagement, immersion, and presence subscales predictors also explained approximately 35% of the variance in overall classification accuracy, $R^2 = .354$, $F(6, 32) = 2.93$, $p = .022$. Interface Quality was the significant contributor, $\beta = .363$, $p = .028$. This relationship also held for the wringing hands cue with approximately 33% of the variance in wringing hands classification accuracy explained by engagement, immersion, and the presence subscales, $R^2 = .334$, $F(6, 32) = 2.68$, $p = .032$. Once again, the Interface Quality subscale was the significant contributor to wringing hands classification accuracy, $\beta = .384$, $p = .022$.

Massed Exposure Group Regressions

In the Massed Exposure condition, the second set of predictor variables, including engagement, immersion, and the presence subscales, revealed significant regression models for

the wringing hands cue only. Engagement, immersion, and the presence subscales explained approximately 35% of the variance in wringing hands detection accuracy, $R^2 = .351$, $F(6, 31) = 2.79$, $p = .027$. It was found that the Interface Quality subscale was the greatest contributor, $\beta = .502$, $p = .003$, and immersion was the secondary contributor, $\beta = .460$, $p = .047$.

Nearly 34% of the variance in the percent change in wringing hands classification accuracy was explained by the engagement, immersion, and the presence subscale predictors, $R^2 = .339$, $F(6, 31) = 2.65$, $p = .034$. The greatest significant contributor was the Natural Interaction subscale, $\beta = -.427$, $p = .035$, and the secondary contributor was the Resolution subscale, $\beta = .400$, $p = .032$.

Highlighting Group Regressions

There were also few significant multiple regression models in the Highlighting condition. The first set of predictor variables indicated that engagement, immersion, and total presence explained nearly 27% of the variance in the percent change in check six detection accuracy, $R^2 = .268$, $F(3, 35) = 4.28$, $p = .011$. Immersion was the greatest significant contributor, $\beta = -.570$, $p = .011$, and total presence was the secondary significant contributor, $\beta = .406$, $p = .017$. Using the second set of predictor variables, it was revealed that engagement, immersion, and the presence subscales explained nearly 32% of the variance in the percent change in check six detection accuracy, $R^2 = .318$, $F(6, 32) = 2.49$, $p = .043$, and indicated immersion as the significant contributor, $\beta = -.613$, $p = .008$.

Finally, engagement, immersion, and total presence also explained approximately 25% of the variance in the percent change in clenched fists median response time, $R^2 = .254$, $F(3, 35) = 3.97$, $p = .015$, with engagement as the significant contributor, $\beta = -.515$, $p = .026$.

CHAPTER FIVE: DISCUSSION

Instructional strategy literature purports Massed Exposure as a viable method to improve attentional weighting and pattern recognition (Carroll, Milham, & Champney, 2009; Williams & Ward, 2003) and Highlighting as a means to improve visual search skills (Chapman, Underwood, & Roberts, 2002), but there is no theoretical nor empirical evidence indicating that one strategy is better than the other. Unfortunately, the lack of significant detection and classification accuracy results did not provide sufficient evidence to assert any advantage of either strategy over the Control for RAISR tasks. However, negative skews of detection and classification accuracy toward higher scores indicate there was a tendency for participants in all three groups to perform well, regardless of training condition. The lack of significant differences and skew toward higher scores may have been driven by the quality of the computer-based content training provided to all participants prior to training scenarios. The content training in the present experiment was aligned with objectives and standards found in current Combat Profiling curriculum and ISR training doctrine (Colombo, Dolletski-Lazar, Coxe, & Tarr, 2012; U.S. Army, 2012a). Therefore, these findings may also serve as a testament to the quality of the content and presentation of information in current behavior cue analysis resources. Further, these results suggest that a computer-based training medium may be sufficient for training IP to accurately detect and classify behavior cues.

The only significant performance results between conditions were revealed for the median response time metric. Sports psychology research indicates that Highlighting may improve reaction time to recognize specific cues (Hagemann, Strauss, & Canal-Bruland, 2006; Jackson & Farrow, 2005); however, in the present study, Massed Exposure, not Highlighting,

yielded response times that were 0.54 seconds faster for the clenched fists cue and 0.37 seconds faster for the slapping hands cue compared to the Control. Given the small effect sizes, the magnitude of impact evident in these findings may seem insufficient to warrant immediate adoption of the Massed Exposure strategy into training systems design requirements for RAISR, thus supporting retention of the null hypothesis. However, variations in response time may impart noticeable consequences on rapid decision-making when attempting to identify HVIs or an impending attack (Lackey & Salcedo, 2014). Although RAISR removes Soldiers and IP from direct line-of-fire (DOD, 2013), as little as half a second may mean the difference in saving the life of an innocent bystander or mitigating the potentially catastrophic impact of an attack. Therefore, Massed Exposure may be a viable SBT strategy to improve response time during behavior cue analysis.

Analyses across conditions revealed a significant effect of cue type on performance. The clenched fists cue had the lowest detection and classification accuracy scores and slowest response time. The slapping hands cue had the highest detection and classification accuracy scores and one of the fastest response times. These cue type effects suggest that the clenched fists cue was the most difficult to distinguish, while the slapping hands cue was easier to identify. Perhaps the clenched fists cue was more subtle because when curling and squeezing the fingers into the palms, the arms are extended downward with the hands held near the body (Givens, 2002). The slapping hands cue may have been more conspicuous because the arms are raised above the waist and the hands are slapped together in front of the body (Lackey, Maraj, Salcedo, Ortiz, & Hudson, 2013).

For percent change performance outcomes, the clenched fists cue had the largest increase in detection accuracy and median response time versus the slapping hands cue which had the

smallest increase in detection accuracy and a decrease in median response time. Notably, the increase in median response time for the clenched fists cue was likely driven by the considerable increase in detection accuracy. The clenched fists cue also had a greater increase in classification accuracy compared to the slapping hands cue, however, the check six cue showed the greatest improvement in classification accuracy. These significantly different changes in performance per cue type suggest that the sample population had varying amounts of prior knowledge about each cue. It appears clenched fists may have been a more novel cue to detect, while the minimal change in performance for the slapping hands cue indicates participants may have been more familiar identifying and interpreting this cue prior to the experiment. Also, the larger increase in classification accuracy for the check six cue indicates participants were likely less familiar with the meaning of this cue prior to the experiment.

High levels of engagement, immersion, and presence have been associated with greater learner satisfaction (Wefald & Downey, 2009; Levett-Jones, et al., 2011; Douglas & Hargadon, 2001; Murphy, 2011; Zhang & Ziguers, 2009; Bulu, 2012). In the present study, perception results revealed very little difference in engagement and immersion and no significant difference in presence between conditions suggesting that the level of satisfaction with the training was consistent across conditions. An analysis of individual items on the Engagement Measure indicated that the Control offered the preferred amount of challenge during training scenarios over the Massed Exposure and Highlighting conditions. Additionally, an analysis of individual items on the Immersion Measure revealed that training scenarios in the Massed Exposure condition were rated equally as challenging as those in the Control, while the Highlighting training scenarios were rated significantly less challenging than the Control. Perhaps participants in the Highlighting group felt the training was too easy and those in the Massed Exposure group

felt the training was more difficult. Despite the plausibility of this explanation, there was only a marginal difference in these item scores between groups, which affirms retention of the null hypothesis for perception measures.

Results of the correlation and regression analyses between training perception and post-test performance were aligned with current VE, simulation, and game-based research claiming that engagement, immersion, and presence during a simulated experience often impact performance (McNamara, Jackson, & Graesser, 2009; Jennett, et al., 2008; Witmer & Singer, 1998). However, the degree to which each perception measure correlated or influenced performance outcomes varied between conditions.

Correlation analysis generally revealed that higher levels of engagement, immersion, and presence were associated with better and improved performance for several performance metrics in the Control. However, these correlations did not hold for the Massed Exposure and Highlighting conditions. In the Massed Exposure condition, improvement in detection and classification accuracy performance were primarily associated with lower ratings of perceived presence. The Highlighting condition revealed a similar relationship with improved clenched fists detection accuracy associated with lower engagement and immersion ratings. These findings suggest that perceived levels of engagement, immersion, and presence during Massed Exposure and Highlighting training exercises imparted little impact on performance outcomes. Therefore, SBT without additional instructional support, as in the Control, may be sufficient for training behavior cue analysis with RAISR. However, the correlations between a reduction in clenched fists response time and higher engagement and immersion in the Highlighting condition point to the potential of the Highlighting condition to improve response times when detecting subtle and novel cues.

Regression analysis results suggest that presence was the predominate predictor of performance. Control group regression analyses revealed that higher perceived presence, specifically higher Involvement and Control and Interface Quality, made a positive contribution to better detection and classification accuracy overall and for select target behavior cues (i.e., clenched fists, check six, and wringing hands). Massed Exposure regressions revealed that higher Interface Quality contributed to better detection accuracy for the wringing hands cue only, but higher Natural Interaction made a negative contribution to wringing hands classification accuracy. In the Highlighting condition, although it was not the primary predictor, presence contributed to improved detection accuracy for the check six cue.

Engagement and immersion appear to have made a negative contribution to performance. Higher engagement did not improve slapping hands detection accuracy in the Control, nor did it improve clenched fists detection accuracy in the Highlighting condition. Furthermore, immersion did not improve check six detection accuracy in the Highlighting condition.

Collectively, results of the correlations and regressions partially supported the alternate hypothesis pertaining to the relationships between performance and perception outcomes with most of the significant relationships revealed in the Control group. Although the difference in presence ratings was not significant between groups, the high frequency of significant, positive correlations and regressions between presence and performance in the Control group suggests that presence, opposed to engagement and immersion, is the more critical subjective aspect of effective SBT for behavior cue analysis during RAISR. Furthermore, the limited and inconsistent presence and performance relationships for Massed Exposure and Highlighting suggest there may be some inherent phenomena of these strategies or the present application hindering the perception of presence during training. In a perceptual skills training context, Massed Exposure

involves increasing the frequency of practice opportunities in a shorter timeframe (Carroll, Milham, & Champney, 2009; Hirumi & Stapleton, 2009) and Highlighting involves orienting the learner's attention to critical content during skill acquisition and practice opportunities using a non-content feature (Carroll, Milham, & Champney, 2009; Chapman, Underwood, & Roberts, 2002). In simulated experiences, presence is characterized as a feeling of being part of the simulated environment to the extent that awareness of the real-world is limited (Slater, Linakis, Usoh, Kooper, & Street, 1996; McMahan, 2003). In the present study, perhaps the higher volume of targets and, consequently, greater difficulty of the Massed Exposure training scenarios induced some degree of distress that prevented participants from feeling fully present, thus causing participants to maintain consciousness of the real-world. Additionally, while the translucent blue box applied in the Highlighting training scenarios seemed to simplify the task, perhaps its depiction in the simulated environment felt unnatural and distracting in the context of the behavior cue analysis task and caused participants to maintain a heightened awareness of the non-content feature, thus inhibiting a feeling of presence. Ultimately, it may not be the strategies themselves that are the issue, but the current format that is preventing a feeling of presence.

CHAPTER SIX: CONCLUSION

The objective of the present experiment was to assess the effectiveness of the Massed Exposure and Highlighting instructional strategies versus a Control for SBT of behavior cue analysis via RAISR. The analyses of participants' performance and perception outcomes provide a foundational evaluation of the relative and potential effectiveness of the strategies. Results of the present study were limited in their capacity to definitively esteem one strategy above the other, however, several unique findings still emerged.

Although Massed Exposure and Highlighting may not contribute to markedly improved detection and classification accuracy nor offer a substantially more positive training experience than standard SBT, neither strategy appears to pose a significant detriment to performance outcomes. However, it was the Massed Exposure strategy that revealed the most compelling utility and extended the very limited instructional design theory regarding its applicability in perceptual tasks. The collective performance results, for both between and across condition analyses, suggest that Massed Exposure has the potential to improve response times when detecting a variety of cues, from the subtle to the conspicuous and from the novel to the familiar. This is evidence of the advantage of Massed Exposure in an ISR context because faster response times contribute to safer operations by enabling early identification of threats. In application, the time and cost to incorporate Massed Exposure into the SBT framework of a PC-based simulation for behavior cue analysis training is minimal; it simply requires the instructional designer or scenario developer to increase the amount of target events when defining the input parameters for the training scenario. Therefore, based on the criticality of response times in ISR and given the performance results and simplicity of implementation, Massed Exposure is the more practical

and accessible SBT strategy recommendation to improve IP ability to detect and classify human behavior more quickly than a standard SBT method.

There were several limitations that may have impeded the ability of the experiment to provide a fully comprehensive assessment of the impact and effectiveness of each strategy on performance and perception. Types of limitations included those related to the content of the experiment and those regarding the analysis of data. These limitations indicate opportunities for further investigation to extend and refine the findings of the present experiment.

One content limitation was that only four target behavior cues from the Kinesics domain were included in the training content and scenarios of the experiment. Since the range of human behavior in the Kinesics domain is extensive, it may be necessary to replicate the assessment conducted in the present study to determine if the results hold for 3D animations developed and validated for other cue types.

Another content limitation involves the means of implementation of the Massed Exposure and Highlighting strategies during training scenarios. The results suggest that the tested strategies did not produce particularly positive training perceptions, which may be attributed to the ratio of target versus non-targets in the Massed Exposure scenarios and the design of the non-content feature in the Highlighting scenarios. This potential limitation presents an opportunity to further investigate the design and delivery of instructional strategies for SBT to improve trainee perceptions. Empirically assessing variations in target to non-target ratios and the design of non-content features will provide a more comprehensive understanding of the utility and optimal delivery methods of each strategy.

A limitation with the analysis of the percent change in classification accuracy variables presents an opportunity to employ an alternate method to calculate the type of performance data

collected during behavior cue analysis or similar tasks. Values for the percent change in classification accuracy variables were driven by the pre-test and post-test detection accuracy scores. In the present experiment, the pre-test and post-test classification accuracy scores used to compute the percent change in classification accuracy were calculated by dividing the number of correctly classified targets by the total number of targets in the scenario. In order to provide a more definitive assessment of the change in classification accuracy, follow-on analyses should adjust the classification accuracy scores to account for the differences in detection accuracy between the pre-test and post-test. This may be achieved by dividing the number of correctly *classified* targets by the number of correctly *detected* targets. The resulting values would be *adjusted* pre-test and post-test classification accuracy scores, which, when used to compute the percent change variables, will result in a value that will be more descriptive of participants' classification ability.

The final limitation concerns both the content and analysis of the experiment. It is possible that the duration of the pre-test and post-test scenarios may have impacted the validity of the performance outcomes. Scenario durations were a consequence of the number of scenario events required to equally balance all the possible model type, cue type, and tetrad position combinations throughout the scenario. Scenario duration may have confounded the impact of the instructional strategies on performance and increased the risk of a type II error. A recommendation for follow-on analyses entails chunking performance data into smaller intervals by time or number of scenario events and assessing behavior cue analysis performance by comparing the differences and changes in accuracy and response time from one interval to the next. An interval analysis may also help to determine the minimum duration or number of training events required to achieve mastery of the behavior cue analysis for each cue type.

Ultimately, the findings of the present experiment derived three implications regarding the theory, application, and further investigation of SBT instructional strategies. The theoretical implication is that the results provide a refined understanding of the utility of the Massed Exposure strategy to improve response time during training of perceptual tasks, which had been previously unassessed. Notably, the present experiment is among the first to identify this unique advantage of the Massed Exposure strategy. Also, the simple implementation of the Massed Exposure strategy by increasing the probability of targets versus non-targets in the present experiment has implications on the instructional design of future SBT for behavior cue analysis with RAISR because it provides an empirically assessed, easily executable, and cost-effective method to apply the Massed Exposure strategy. Finally, empirical implications resulting from the limitations of the present experiment petition the instructional design community to respond with viable solutions to address these gaps and continually expand the body of work related to SBT instructional strategies for training perceptual tasks.

APPENDIX A: ENGAGEMENT MEASURE

Engagement Measure

INSTRUCTIONS

For each statement, select the number that indicates how much you agree or disagree with the statement.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

1. It would not matter to me if I never use a virtual environment for behavior cue detection training in again.
2. I felt happy at the thought of using a virtual environment for behavior cue detection training.
3. The less I have to do with using a virtual environment for behavior cue detection training, the better.
4. I want to spend more time using a virtual environment for behavior cue detection training.
5. It was important to me to be good at behavior cue detection.
6. I experienced a “buzz of excitement” while using the virtual environment for behavior cue detection training.
7. I like the challenge that using a virtual environment for behavior cue detection training provided.

SCORING

Items 1 and 3 are reverse scored. Sum the responses for all items to derive the total Engagement Measure score.

Adapted from: Charlton, J. & Danforth, I. (2005). Distinguishing addiction and high engagement in the context of online game playing. Computers in Human Behavior, 23(3), 1531-1548.

APPENDIX B: IMMERSION MEASURE

Immersion Measure

INSTRUCTIONS

For each statement, select the number that indicates how much you agree or disagree with the statement.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly Agree

1. I was interested in seeing how the scenario events would progress.
2. I was in suspense about whether I would perform well or not in the scenarios.
3. I sometimes found myself to become so involved with the scenarios that I wanted to speak to the scenarios directly.
4. I enjoyed the graphics and imagery of the scenarios.
5. I enjoyed completing the scenarios.
6. I was unaware of what was happening around me.
7. I feel that I tried my best during the scenarios.
8. The scenarios were challenging.

SCORING

Sum the responses for all items to derive the total Immersion Measure score.

Adapted from: Jennett, C., Cox, A., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9), 641-661.

APPENDIX C: PRESENCE MEASURE

Presence Questionnaire

INSTRUCTIONS

Select a value on the scale in accordance with the question content and descriptive labels.

- | | | | | | | | |
|---|-----------------------|---|---|---------|---|---|-----------------|
| 1. How much were you able to control events? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all | | | Neutral | | | Very much |
| 2. How responsive was the environment to actions that you initiated (or performed)? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all responsive | | | Neutral | | | Very responsive |
| 3. How natural did your interactions with the environment seem? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all natural | | | Neutral | | | Very natural |
| 4. How much did the visual aspects of the environment involve you? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all | | | Neutral | | | Very much |
| 5. How natural was the mechanism that controlled movement through the environment? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all natural | | | Neutral | | | Very natural |
| 6. How compelling was your sense of objects moving through space? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all compelling | | | Neutral | | | Very compelling |
| 7. How much did your experiences in the virtual environment seem consistent with your real world experiences? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all | | | Neutral | | | Very much |
| 8. Were you able to anticipate what would happen next in response to the actions that you performed? | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | Not at all | | | Neutral | | | Very much |

9. How completely were you able to actively survey or search the environment using vision?

1	2	3	4	5	6	7
Not well			Neutral			Very well

10. How compelling was your sense of moving around inside the virtual environment?

1	2	3	4	5	6	7
Not at all compelling			Neutral			Very compelling

11. How closely were you able to examine objects?

1	2	3	4	5	6	7
Not at all closely			Neutral			Very closely

12. How well could you examine objects from multiple viewpoints?

1	2	3	4	5	6	7
Not well			Neutral			Very well

13. How involved were you in the virtual environment experience?

1	2	3	4	5	6	7
Not at all involved			Neutral			Very involved

14. How much delay did you experience between your actions and expected outcomes?

1	2	3	4	5	6	7
Not at all			Neutral			Very much

15. How quickly did you adjust to the virtual environment experience?

1	2	3	4	5	6	7
Not at all			Neutral			Very quickly

16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

1	2	3	4	5	6	7
Not at all			Neutral			Very proficient

17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

1	2	3	4	5	6	7
Not at all			Neutral			Very much

18. How much did the control devices interfere with the performance of assigned tasks or with other activities?

1	2	3	4	5	6	7
Not at all			Neutral			Very much

19. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

1	2	3	4	5	6	7
Not well			Neutral			Very well

20. How easily did you adjust to the control devices used to interact with the virtual environment?

1	2	3	4	5	6	7
Not at all			Neutral			Very easily

SCORING

For each subscale, sum the responses for the items listed. Note that items 17 and 18 are reverse scored.

- Involvement/Control: Items 1, 2, 4, 6, 8, 9, 10, 13, 14, 15, 16, 20
- Natural Interaction: Items 3, 5, 7
- Resolution: Items 11, 12
- Interface Quality: Items 17r, 18r, 19

APPENDIX D: INFORMED CONSENT FORM



STRIVE 2: Systems Training Research in Virtual Environments 2

Informed Consent

Principal Investigator: Stephanie Lackey, Ph.D.
Co-Investigator: Julie Salcedo
Sponsor: ARL – U.S. Army Research Laboratory
Investigational Site: Institute for Simulation and Training
University of Central Florida
3100 Technology Parkway, Orlando, FL 32826

Introduction: Researchers at the University of Central Florida (UCF) study many topics. To do this we need the help of people who agree to take part in research studies. You are being invited to take part in a research study which will include up to 150 people at UCF. You must be 18 to 40 of age, have U.S. citizenship, have normal or corrected to normal vision, and must not be colorblind to be included in the research study. Your participation is voluntary and you may withdraw at any time. The people conducting this research are Stephanie Lackey and Julie Salcedo from the Institute for Simulation and Training at UCF.

What you should know about a research study:

- Someone will explain this research study to you.
- A research study is something you volunteer for.
- Whether or not you take part is up to you.
- You should take part in this study only because you want to.
- You can choose not to take part in the research study.
- You can agree to take part now and later change your mind.
- Whatever you decide it will not be held against you.
- Feel free to ask all the questions you want before you decide.

Purpose of the research study: The purpose of this research is to understand the application of Unmanned Ground Systems (UGSs) for Intelligence, Surveillance, and Reconnaissance (ISR) missions. The focus is on virtual training technologies and strategies for Tactical Operations Center operators to monitor, prioritize, and coordinate UGS video and communication sources. The objective of this experiment is to compare the effects between various Simulation-Based Training strategies for a human behavior analysis task.

What you will be asked to do in the study: You will view instructional PowerPoint presentations demonstrating how to detect and classify various human behavioral cues from the perspective of an UGS. You will complete several training scenarios within a simulated

environment presented on a desktop computer monitor. You will use a computer mouse to practice the detection and classification tasks by selecting targets. You will be asked to complete a variety of questionnaires throughout the experiment including: demographics, immersive tendencies, current health status, and your perceptions of immersion, engagement, flow, presence, technology acceptance, intrinsic motivation, and workload.

Location: This study is being conducted in the ACTIVE Lab (Rooms 306H or 337) at the Institute for Simulation and Training: 3100 Technology Parkway Orlando, FL 32826.

Time required: The expected duration of this study will not exceed 5 hours.

Funding for this study: This research study is funded by the U.S. Army Research Laboratory (ARL).

Risks: There are no foreseeable risks or discomforts other than those normally encountered in the daily lives of healthy persons. There is minimal risk that you may develop what is referred to as simulator sickness. It periodically occurs after exposure to prolonged, continuous testing in simulated environments. Symptoms consist of nausea, disorientation, and a visual disruption. The Simulator Sickness Questionnaire will be used to monitor you for symptoms. The risk is minimized as a result of the short duration of each session within the simulated environment. If you experience any of the symptoms mentioned, please tell the researcher and remain seated until the symptoms subside.

Benefits: We cannot promise any benefits to you or others from your taking part in this research. However, possible educational benefits include a better understanding of human behavior.

Compensation: You will be compensated \$10 per hour of the session OR course credit at the discretion of your professor. For paid participants, a receipt of completion will be provided at the end of the session which may be redeemed at the IST SONA cashier located on the 3rd floor at 3100 Technology Parkway, Orlando, FL 32826. For credit participants, credit will be awarded on the IST SONA System upon completion of the study and will be applied as course credit at the discretion of your course professor. If you choose not to participate in the study at this time, then you will not receive compensation. If you consent to participate, but later withdraw or must be dismissed, then you will be compensated for the amount of time you participated.

Confidentiality: We will limit access to data collected in this study to people who have a need to review this information. We cannot promise complete secrecy. The principal investigators, co-investigators, and research assistants working on this project will have access to your data. Additionally, there is a possibility that the U.S. Army Human Research Protections Office (AHRPO) will also review the records related to this study. Data will be secured in locked cabinets at the Institute for Simulation and Training (IST) following UCF IRB protocol. Please note that your name will not be associated with any of the data collected during this study. Once you sign this Informed Consent document, it will be kept in a locked cabinet separate from your data.

Study contact for questions about the study or to report a problem: If you have questions, concerns, complaints, or think the research had a negative impact on your well-being, contact the Principal Investigator, Stephanie Lackey at slackey@ist.ucf.edu or 407-882-2427 or the Co-Investigator, Julie Salcedo, at jsalcedo@ist.ucf.edu or 407-882-0037.

IRB contact about your rights in the study or to report a complaint: Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (UCF IRB). This research has been reviewed and approved by the IRB. For information about the rights of people who take part in research, please contact: Institutional Review Board, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901. You may also talk to them for any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You want to get information or provide input about this research.

Once all of your questions about the study have been answered and if you want to continue your participation in this study please sign below.

The researcher will then take this entire informed consent and place it in a locked cabinet separate from your data. You will be given another copy of the exact same informed consent for you to keep.

Participant printed name

Participant signature

Date

Printed name of person obtaining consent

Signature of person obtaining consent

Date

APPENDIX E: DEMOGRAPHICS QUESTIONNAIRE

Demographics Questionnaire

1. Age: _____
2. Sex: (Circle one) FEMALE MALE
3. Which is your predominate hand? (Circle one) RIGHT LEFT
4. Are you color blind? (Circle one) YES NO
5. Do you have normal/corrected vision? (Circle one) YES NO
If YES, are you wearing corrected lenses now? (Circle one) YES NO
6. Are you in your usual state of health? (Circle one) YES NO
If NO, briefly explain: _____
7. Approximately, how many hours of sleep did you get last night? _____
8. What is your major? (If applicable) _____
9. Have you ever served in the military or ROTC? (Circle one) YES NO
If YES, when and/or what branch? _____
10. What is your occupation? _____
11. What is your highest level of education completed? (Circle one)
High School or equivalent Less than 4 yrs of college Completed 4 yrs of college
More than 4 yrs of college Other: _____
12. When did you use computers in your education? (Circle all that apply)
Grade School Jr. High High School Technical School College
Did Not Use
13. Please estimate the number of hours you use a computer per week (If none, write "0"): _____

14. Where do you currently use a computer? (Circle all that apply)

Home Work Library Other: _____ Do Not Use

15. How would you describe your degree of comfort with computer use? (Circle one)

Poor Fair Average Above average Proficient

16. For each of the following questions, circle the response that best describes you.

How often do you:

Use a mouse? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use a joystick? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use a keyboard? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use a touchscreen? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use a game controller? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use icon-based program/software? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use graphics/drawing features in software programs?

Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use email? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Operate radio/remote controlled vehicle/device (e.g., RC car, boat, or plane)?

Daily, Weekly, Monthly, Once every few months, Rarely, Never

Play computer/video games? Daily, Weekly, Monthly, Once every few months, Rarely, Never

17. Please estimate the number of hours you play video games per week (If none, write "0"): _____

If you play video games, which types of games do you play? (Circle all that apply)

First-person shooters Strategy Sports Racing Other: _____

APPENDIX F: EXAMPLES OF TARGET BEHAVIOR CUE MODELS

Slap Hands (Aggressiveness)



Clench Fists (Aggressiveness)



Wring Hands (Nervousness)



Check Six (Nervousness)



APPENDIX G: KINESIC CUE REVIEW SCORE SHEET

Kinesic Cue Review Score Sheet

INSTRUCTIONS

Circle the participant's response. For Other, write in the response.

- | | | | | |
|----|----------------|-------------|------------|-------------|
| 1. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 2. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 3. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 4. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 5. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 6. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 7. | Aggressiveness | Nervousness | Don't Know | Other _____ |
| 8. | Aggressiveness | Nervousness | Don't Know | Other _____ |

APPENDIX H: INSTITUTIONAL REVIEW BOARD APPROVAL LETTER



University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Stephanie Jane Lackey and Co-PI: Julie N. Salcedo**

Date: **January 14, 2014**

Dear Researcher:

On 1/14/2014, the IRB approved the following modifications / human participant research until 06/11/2014 inclusive:

Type of Review: IRB Addendum and Modification Request Form
Expedited Review Category #7

Modification Type: The following research assistants are being added to the study:
Veronica Prisco, Heidi Buehrer, and Eric Ortiz

Project Title: STRIVE 2: Systems Training Research in Virtual Environments
2

Investigator: Stephanie Jane Lackey

IRB Number: SBE-13-09424

Funding Agency: US Army Research Laboratory

Grant Title:

Research ID: 1052585

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30 days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form cannot be used to extend the approval period of a study. All forms may be completed and submitted online at <https://iris.research.ucf.edu>.

If continuing review approval is not granted before the expiration date of 06/11/2014, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a signed and dated copy of the consent form(s).

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

IRB Coordinator

LIST OF REFERENCES

- Abich, J., Taylor, G., & Reinerman-Jones, L. (2013). Establishing workload manipulations utilizing a simulated environment. *Proceedings of the 8th International Conference of HCI International*. Las Vegas, NV.
- Army Research Laboratory. (2011). *Robotics Collaborative Technology Alliance (RCTA) FY 2011 Annual Program Plan*. ARL.
- Arthur, Jr., W., Bennett, Jr., W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, 88(2), 234-245.
- Bardu, A., Bridge, A., Coroian, D., Dickinson, S., Mussman, S., Narayanaswamy, S., . . . Zhang, Z. (2012). Large-scale automatic labeling of video events with verbs based on event-participant interaction. *Computing Research Repository*. Retrieved from <http://arxiv.org/abs/1204.3616>
- Becker, W., & Schatz, S. (2010). Higher Order Cognitive Skills Training in Support of Contemporary Military Operations. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2228-2232). Santa Monica, CA: Sage.
- Benenson, J. F., Nicholson, C., Waite, A., Roy, R., & Simpson, A. (2001). The Influence of Group Size on Children's Competitive Behavior. *Child Development*, 72(3), 921-928.
- Birdwhistell, R. L. (1970). *Kinesics and context: Essays on body motion communication*. Philadelphia, PA: University of Pennsylvania Press.
- Bohemia Interactive. (2014). *VBS2*. Retrieved from Bohemia Interactive Simulations: <http://products.bisimulations.com/products/vbs2/overview>

- Bouma, H., Hanckmann, P., Marck, J.-W., Penning, L., den Hollander, R., ten Hove, J.-M., . . . Burghouts, G. (2012). Automatic human action recognition in a scene from visual inputs. *Proc. SPIE 8388, Unattended Ground, Sea, and Air Sensor Technologies and Applications XIV*.
- Bowman, D. A., & McMahan, R. P. (2007, July). Virtual reality: How much immersion is enough? *Computer*, pp. 36-43.
- Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual Perception* (Fourth ed.). New York, NY: Taylor and Francis.
- Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers and Education*, 58(1), 154-161.
- Carroll, M., Milham, L., & Champney, R. (2009). Military observations: Perceptual skills training strategies. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)* (p. No. 9287). Arlington, VA: NDIA.
- Chapman, P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Psychology and Behaviour*, 5, 157-167.
- Charlton, J., & Danforth, I. (2005). Distinguishing addiction and high engagement in the context of online game playing. *Computers in Human Behavior*, 23(3), 1531-1548.
- Chen, J. Y., Quinn, S., Wright, J., Barnes, M., Barber, D., & Adams, D. (2013). Human-agent teaming for robot management in multitasking environments. *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, (pp. 103-104). Tokyo, Japan.

- Coba, J. V. (2010). Application OF Copper Indium Gallium Diselenide Photovoltaic Cells to Extend the Endurance and Capabilities of the Raven RQ-11B Unmanned Aerial Vehicle. Naval Postgraduate School.
- Colombo, G., Dolletski-Lazar, R., Coxe, M., & Tarr, R. (2012). Setting the stage: Preparation for advanced Combat Profiling training. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2012*, (p. 12063).
- Coover, J. E. (1913). The feeling of being stared at. *The American Journal of Psychology*, 24(4), 570-575.
- Csikszentmihalyi, M. (1990). *Flow, the Psychology of Optimal Experience*. New York: Harper Collins.
- Dalzell, T. (2009). *The Routledge Dictionary of Modern American Slang and Unconventional English*. New York: Routledge.
- DARPA. (2011, January 4). *DARPA kicks off Mind's Eye program*. Retrieved from DARPA: http://www.darpa.mil/NewsEvents/Releases/2011/2011/01/04_DARPA_Kicks_Off_Mind's_Eye_Program.aspx.
- de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology*, 21, 731-746.
- de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20(2), 111-122.
- de Penning, H., den Hollander, R., Bouma, H., Burghouts, G., & d'Avila Garcez, A. (2012). *A neural-symbolic cognitive agent with a Mind's Eye*. Association for the Advancement of Artificial Intelligence.

- DOD. (2013). *Unmanned Systems Integrated Roadmap FY2013-2038*. Washington D.C.: Department of Defense.
- Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, 84(5), 795-805.
- Douglas, J. Y., & Hargadon, A. (2001). The pleasures of immersion and engagement: Schemas, scripts, and the fifth business. *Digital Creativity*, 12(3), 153-166.
- Dunne, R., Schatz, S., Fiore, S. M., Martin, G., & Nicholson, D. (2010). Scenario-Based Training: Scenario Complexity. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2238-2242). Santa Monica, CA: Sage.
- Dunne, R., Schatz, S., Fiore, S. M., Nicholson, D., & Fowlkes, J. (2010). *Optimizing Decision Preparedness by Adapting Scenario Complexity and Automating Scenario Generation*. Hampton, VA: NASA Langley Research Center.
- Eastin, M. S. (2007). The Influence of Competitive and Cooperative Group Game Play on State Hostility. *Human Communication Research*, 33(4), 450-466.
- Eckstein, M. P., Abbey, C. K., Pham, B. T., & Shimozaki, S. S. (2004). Perceptual learning through optimization of attentional weighting: Human versus optimal Bayesian learner. *Journal of Vision*, 4(12), 1006-1019.
- Fautua, D., & Schatz, S. (2010). *Border Hunter Research Technical Report*. Norfolk, VA: U.S. Joint Forces Command.
- Fiore, S. M., Scielzo, S., Jentsch, F., & Howard, M. L. (2006). Understanding performance and cognitive efficiency when training for x-ray security screening. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 2610-2614). Santa Monica, CA: Sage.

- Fiorella, L., Vogel-Walcutt, J. J., & Fiore, S. (2012). Differential impact of two types of metacognitive prompting provided during simulation-based training. *Computers in Human Behavior*, 696-702.
- Fletcher, J. D. (2004). *Cognitive Readiness: Preparing for the Unexpected*. Institute for Defense Analyses. Alexandria, VA: DOD.
- Flynn, M. (2010). Left of bang: The value of socio-cultural analysis in today's environment. In A. Astorino-Courtois, H. Cabayan, B. Casebeer, A. Chapman, D. DiEuliis, C. Ehlschlaeger, . . . C. Rice, *National Security Challenges: Insights from Social, Neurobiological, and Complexity Sciences* (pp. 1-11). Washington, D.C.: U.S. Army ERDC.
- Fowlkes, J., Dwyer, D. J., Oser, R. L., & Salas, E. (1998). Event-Based Approach to Training. *The International Journal of Aviation Psychology*, 8(3), 209-221.
- Fowlkes, J., Norman, J. W., Schatz, S., & Stagl, K. C. (2009). Contrasting Cases: A Strategy for Advanced Learning Using Simulation-based Training. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp. 1935-1938). Santa Monica, CA: Sage.
- Fuchs, T., McNevin, N. H., Ritter, F., Toole, T., & Wulf, G. (2000). Attentional focus in complex skill learning. *Research Quarterly for Exercise and Sport*, 71(3), 229-239.
- Gideons, C. D., Padilla, F. M., & Lethin, C. R. (2008, September). Combat Hunter: The training continues. *Marine Corps Gazette*, 92(9), 79-84.
- Givens, D. B. (2002). *The Nonverbal Dictionary of Gestures, Signs, and Body Language Cues*. Spokane, WA: Center for Nonverbal Studies Press.

- Hagemann, N., Strauss, B., & Canal-Bruland, R. (2006). Training perceptual skill by orienting visual attention. *Journal of Sport and Exercise Psychology*, 28, 143-158.
- Hale, K., Carpenter, A., Johnston, M., Costello, J.-J., Flint, J., & Fiore, S. M. (2012). Adaptive training for visual search. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2012*, (p. 12144).
- Havice, P. A., Davis, T. T., Foxx, K. W., & Havice, W. L. (2010). The impact of rich media presentations on a distributed learning environment: Engagement and satisfaction of undergraduate students. *The Quarterly Review of Distance Education*, 11(1), 53-58.
- Hirumi, A., & Stapleton, C. (2009). Applying pedagogy during game development to enhance game-based learning. In C. Miller, *Games: Purpose and Potential in Education* (pp. 127-162). New York, NY: Springer.
- Ishihara, S. (2013). *Ishihara's Tests for Colour Deficiency*. Tokyo, Japan: Kanehara Trading.
- Jackson, R. C., & Farrow, D. (2005). Implicit perceptual training: How, when, and why? *Human Movement Science*, 24(3), 308-325.
- Jennett, C., Cox, A., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9), 641-661.
- Kass, S. J., Herschler, D. A., & Companion, M. A. (1991). Training situational awareness through pattern recognition in a battlefield environment. *Military Psychology*, 3(2), 105-112.
- Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society of Training and Development*, 13, 3-9.

- Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig, *Training and Development Handbook: A Guide to Human Resource Development* (2nd ed., pp. 301-319). New York: McGraw-Hill.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating Training Programs: The Four Levels* (3rd ed.). San Francisco, CA: Berrett-Koehler Publishers, Inc.
- Kotnour, T., Landaeta, R., & Lackey, S. (2013). Defining training system impact assessment measures from a stakeholder perspective: Case study of the NEW-IT project. *International Journal of Technology Enhanced Learning*, 5(1), 1-23.
- Lackey, S. J., & Salcedo, J. N. (2014). Assessing instructional strategies for training robot-aided ISR tasks in simulated environments. *Proceedings of the Spring Simulation Multiconference, Annual Simulation Symposium*. Tampa, FL.
- Lackey, S. J., Maraj, C., & Barber, D. J. (2014). Immersion, presence, and flow in robot-aided ISR simulation-based training. *Proceedings of the Winter Simulation Conference*. Savannah, GA.
- Lackey, S. J., Maraj, C., Salcedo, J., Ortiz, E., & Hudson, I. (2013). Assessing performance of kinesic cue analysis in simulation-based training environments. *Proceeding of the Interservice/Industry Training, Education, and Simulation Conference (I/ITSEC 2013)*.
- Levett-Jones, T., McCoy, M., Lapkin, S., Noble, D., Hoffman, K., Dempsey, J., . . . Roche, J. (2011). The development and psychometric testing of the satisfaction with simulation experience scale. *Nurse Education Today*, 31(7), 705-710.
- Lin, C.-P. (2009). Learning online brand personality and satisfaction: The moderating effects of gaming engagement. *International Journal of Human-Computer Interaction*, 25(3), 220-236.

- Lyons, D., Schmorow, D., Cohn, J., & Lackey, S. (2002). Scenario based training with virtual technologies and environments. *Proceedings of the IMAGE 2002 Conference*. Scottsdale, AZ.
- MARCBot. (2010). Retrieved from Exponent Engineering and Scientific Consulting:
http://www.exponent.com/marcbot_product/
- Martin, G. A., Hughes, C. E., Schatz, S., & Nicholson, D. (2010). The Use of Functional L-Systems for Scenario Generation in Serious Games. *Proceedings of the Workshop on Procedural Content Generation in Games* (p. No. 6). New York, NY: ACM.
- Martin, G. A., Schatz, S., Bowers, C., Hughes, C. E., Fowlkes, J., & Nicholson, D. (2009). Automatic Scenario Generation Through Procedural Modeling for Scenario-Based Training. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp. 1949-1953). Santa Monica, CA: Sage.
- Martin, G. A., Schatz, S., Hughes, C. E., & Nicholson, D. (2010). What is a Scenario? Operationalizing Training Scenarios for Automatic Generation. In D. Kaber, & G. Boy, *Advances in Cognitive Ergonomics* (pp. 746-753). Boca Raton, FL: CRC Press.
- McMahan, A. (2003). Immersion, engagement, and presence: A method for analyzing 3-D video games. In M. J. Wolf, & B. Perron, *The Video Game, Theory Reader* (pp. 67-86). New York: Taylor & Francis Group.
- McNamara, D. S., Jackson, G. T., & Graesser, A. (2009). Intelligent tutoring and games (ITaG). *Workshop Proceedings of the 14th International Conference on Artificial Intelligence in Education (AIED)*, (pp. 1-10). Brighton, UK.
- Milburn, D. (2012). *Army PM UAS Spectrum Update*. Presented at the Pacific Spectrum Management Conference, Sigmatech, Inc., U.S. Army Unmanned Aircraft Systems.

- MIT. (2010). *The Mind of the Mind's Eye*. Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory.
- Mogg, K., & Bradley, B. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition & Emotion*, 13(6), 713-740.
- Mogg, K., & Bradley, B. (2002). Selective orienting of attention to masked threat faces in social anxiety. *Behavior Research and Therapy*, 40(12), 1403-1414.
- Murphy, C. (2011). Why games work: The science of learning (Tutorial). *Interservice/Industry Training, Simulation, and Education Conference*. Orlando, FL.
- Mykoniatis, K., Angelopoulou, A., Soyler, A., Kincaid, P., & Hancock, P. (2012). Modeling and simulation of Human robot interaction: Face recognition implementations. *IEEE Intelligent Systems*.
- Navarro, J., & Karlins, M. (2008). *What Every Body is Saying*. New York: Harper-Collins.
- Neisser, U. (1964). Visual search. *Scientific American*, 210(6), 94-102.
- Nicholson, D., & Schatz, S. (2010). Making the Implicit Explicit: Operationalizing Tacit Facets of Scenario-Based Training. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2226-2227). Santa Monica, CA: Sage.
- Nicholson, D., Fiore, S., Vogel-Walcutt, J. J., & Schatz, S. (2009). Advancing the science of training in simulation-based training. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp. 1932-1934). Santa Monica, CA: Sage.
- Nicholson, D., Schatz, S., Stanney, K., & Lackey, S. (2009). *Next-generation Expeditionary Warfare Intelligent Training (NEW-IT) Program Summary Booklet*. University of Central Florida, Institute for Simulation and Training. Arlington, VA: Office of Naval Research.

- Nodine, C. F., Mello-Thomas, C., Kundel, H. L., & Weinstein, S. P. (2002). Time course of perception and decision-making during mammographic interpretation. *American Journal of Roentgenology*, 179(4), 917-923.
- Orford, J. (1986). The rules of interpersonal complementarity: Does hostility beget hostility and dominance, submission? *Psychological Review*, 93(3), 365-377.
- Ortiz, E., Maraj, C., Salcedo, J., Lackey, S., & Hudson, I. (2013). Assessing engagement in simulation-based training systems for virtual kinesic cue detection training. *Proceedings of the Human Computer Interaction International Conference (HCII)*. Las Vegas, NV.
- Oser, R. L., Cannon-Bowers, J. A., Salas, E., & Dwyer, D. J. (1999). Enhancing Human Performance in Technology-Rich Environments: Guidelines for Scenario-Based Training. *Human/Technology Interaction in Complex Systems*, 9, 175-202.
- Oser, R. L., Gualtieri, J. W., Cannon-Bowers, J. A., & Salas, E. (1999). Training Team Problem Solving Skills: An Event-Based Approach. *Computers in Human Behavior*, 15, 441-462.
- Pearson, J. E., Moore, Z., Ogdock, J. S., & Choi, F. J. (2013). Design of sense-and-avoid standards for RQ-7B Shadow under loss-link. *2013 IEEE Systems and Information Engineering Design Symposium*. Charlottesville, VA.
- Potegal, M., & Knutson, J. F. (1994). *The Dynamics of Aggression: Biological and Social Processes in Dyads and Groups*. Hillsdale, NJ: Lawrence Erlbaum.
- Poulter, D. R., Jackson, R. C., & Berry, D. C. (2005). The effect of learning condition on perceptual anticipation, awareness, and visual search. *Human Movement Science*, 24(3), 345-361.
- Rezec, A. A., & Dobkins, K. R. (2004). Attentional weighting: A possible account of visual field asymmetries in visual search? *Spatial Vision*, 17(4-5), 269-293.

- Ross, W. A., & Militello, L. G. (2013). Methodology for developing the expert model: The case of the Combat Hunter. In D. D. Schmorrow, & D. M. Nicholson, *Advances in Design for Cross-Cultural Activities: Part I* (pp. 13-22). Boca Raton, FL: CRC Press.
- Ross, W. A., Bencaz, N., & Militello, L. G. (2010). *Specification and development of an expert model for "Combat Hunters"*. U.S. Joint Forces Command.
- Sachdeva, S. (2009). Fitzpatrick skin typing: Applications in dermatology. *Indian Journal of Dermatology, Venereology, and Leprology*, 75(1), 93-96.
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52, 471-499.
- Salcedo, J. N., Lackey, S. J., & Maraj, C. (2014). Impact of instructional strategies on motivation and engagement for simulation-based training of robot-aided ISR tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. San Diego, CA: Sage.
- Salcedo, J. N., Lackey, S. J., & Reinerman-Jones, L. E. (2014). Massed exposure improves response time for detection of nervous behaviors in robot-aided ISR missions. *Proceedings of the MODSIM World Conference*. Society for Modeling and Simulation International.
- Salcedo, J., Maraj, C., Lackey, S., Ortiz, E., Hudson, I., & Martinez, J. (2013). Effects of visual fidelity on biometric cue detection in virtual Combat Profiling training. *Proceedings of the Human Computer Interaction International Conference 2013 (HCII 2013)*. Las Vegas, NV.
- Sanders, D., & Welk, D. S. (2005). Strategies to scaffold student learning: Applying Vygotsky's zone of proximal development. *Nurse Educator*, 30(5), 203-207.

- Schatz, S., Folsom-Kovarik, J. T., Bartlett, K., Wray, R. E., & Solina, D. (2012). Archetypal patterns of life for military training simulations. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2012*, (p. 12193).
- Schatz, S., Reitz, E. A., Nicholson, D., & Fautua, D. (2010). Expanding Combat Hunter: The science and metrics of Border Hunter. *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* (p. 10147). Arlington, VA: NTSA.
- Schumann, P. L., Anderson, P. H., Scott, T. W., & Lawton, L. (2001). A framework for evaluating simulations as educational tools. *Developments in Business Simulation and Experiential Learning*, 28, 215-220.
- Scott, G. (2008). *University Student Engagement and Satisfaction with Learning and Teaching*. University of Western Sydney, Sydney, Australia.
- Seitz, A. R., & Dinse, H. R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology*, 17(2), 148-153.
- Slater, M., Linakis, V., Usoh, M., Kooper, R., & Street, G. (1996). Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. *ACM Virtual Reality Software and Technology*, 163-172.
- Spiker, A., Williams, G., Johnston, J., & Lethin, C. (2010). Training tactical behavior profiling skills for irregular warfare. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, (p. 10118).
- Swanson, R. A., & Sleezer, C. M. (1987). Training effectiveness evaluation. *Journal of European Industrial Training*, 11(4), 7-16.
- Tarr, M. J. (2000). Pattern recognition. In A. E. Kazdin, *Encyclopedia of Psychology* (Vol. 6, pp. 66-71). Washington, DC, US: American Psychological Association.

- U.S. Army. (2007). *FM 2-91.6: Soldier Surveillance and Reconnaissance: Fundamentals of Tactical Information Collection*. Washington, D.C.: Department of the Army.
- U.S. Army. (2008). *TRADOC Pamphlet 525-7-9: The United States Army's Concept Capability Plan (CCP): Intelligence, Surveillance, and Reconnaissance*. Washington, D.C.: Department of the Army.
- U.S. Army. (2009). *FM 2-01: ISR Synchronization*. Washington, D.C.: Department of the Army.
- U.S. Army. (2010). *FM 2-0: Intelligence*. Washington, D.C.: Department of the Army.
- U.S. Army. (2012a). *ADRP 2-0: Intelligence*. Washington, D.C.: Department of the Army.
- U.S. Army. (2012b). *FM 3-55: Information Collection*. Washington, D.C.: Department of the Army.
- Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics*, 50(8), 1235-1249.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*.
- Van den Berg, M. H., Ronday, H. K., Peeters, A. J., Voogt-van der Harst, E. M., Munneke, M., Beedveld, F. C., & Vliet Vlieland, T. M. (2007). Engagement and satisfaction with an Internet-based physical activity intervention in patients with rheumatoid arthritis. *Rheumatology*, 46(3), 545-552.
- Vogel-Walcutt, J. J., Marshall, R., Schatz, S., Dolletski-Lazar, R., & Nicholson, D. (2011). Effects of instructional strategies within an instructional support system to improve knowledge acquisition and application. *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (pp. 2133-2137). Santa Monica, CA: Sage.
- Voth, D. (2004). A new generation of military robots. *Intelligent Systems, IEEE*, 19(4), 2-3.

- Vygotsky, L. (1978). *Mind and Society: The Development of Higher Mental Processes*. Cambridge: Cambridge University Press.
- Wefald, A. J., & Downey, R. G. (2009). Construct dimensionality of engagement and its relation with satisfaction. *The Journal of Psychology*, 143(1), 91-111.
- Williams, A. M., & Ward, P. (2003). Perceptual expertise: Development in sport. In J. L. Starkes, & K. A. Ericsson, *Expert Performance in Sports: Advances in Research on Sport Expertise* (pp. 219-250). Champaign, IL: Human Kinetics.
- Witmer, B., & Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3), 225-240.
- Yantis, S. (2001). *Visual Perception*. Philadelphia, PA: Taylor and Francis.
- Zhang, C., & Ziguers, I. (2009). An exploratory study of the impact of a virtual world learning environment on student interaction and learning satisfaction. *Proceedings of the 15th Americas Conference on Information Systems*, (p. 424).